REPORT RESUMES

ED 002 915 24

PSYCHOLOGICAL PROBLEMS AND RESEARCH METHODS IN MATHEMATICS TRAINING.

BY- DU BOIS, PHILIP H. FEIERABEND, ROSALIND L.
WASHINGTON UNIV., ST.LOUIS, MO.

REPORT NUMBER CRP-642 PUB DATE JUN 59
EDRS PRICE MF-$0.36 HC-$8.84 221P.

DESCRIPTORS- *MATHEMATICS INSTRUCTION, MATHEMATICAL EXPERIENCE, SURVEYS, CONFERENCES, PSYCHOEDUCATIONAL PROCESSES, *RESEARCH METHODOLOGY, *LEARNING PROCESSES, *CURRICULUM GUIDES, MANUALS, TEACHING TECHNIQUES, ST. LOUIS, MISSOURI

A CRITICAL REVIEW OF PSYCHOLOGICAL LITERATURE AND A RESEARCH METHODS INVENTORY, RELATING TO THE LEARNING OF MATHEMATICS AT ALL ACADEMIC LEVELS, WERE CONDUCTED TO DEVELOP A HANDBOOK OR GUIDE FOR USE BY INVESTIGATORS IN THE FIELD OF MATHEMATICS TEACHING. THIS EFFORT TOOK PLACE DURING A 3-MONTH PERIOD WHICH INCLUDED SYSTEMATIC SURVEYS OF LITERATURE AND ONGOING RESEARCH STUDIES, AS WELL AS A COMPREHENSIVE REVIEW WITH VARIOUS CONSULTANTS IN THE FIELD TO OBTAIN INFORMATION ON POSSIBLE APPLICATIONS OF SUCH SPECIALIZED DEVELOPMENTS AS INFORMATION THEORY AND MULTIVARIATE AND FACTOR ANALYSIS IN MATHEMATICS TRAINING. ALL FINDINGS WERE INCORPORATED INTO THE FINAL REPORT WHICH WAS DIVIDED INTO THE FOLLOWING THREE SECTIONS--(1) "PROBLEM AREAS IN THE TEACHING OF MATHEMATICS," (2) "PROPOSALS FOR RESEARCH ON THE TEACHING OF MATHEMATICS," AND (3) "REVIEW OF RESEARCH ON MATHEMATICAL EDUCATION." PROBLEM AREAS COVERED HUMAN LEARNING WITH SPECIAL REFERENCE TO TEACHING OF MATHEMATICS AND INCLUDED MEASUREMENT PROBLEMS, INDUCTIVE AND DEDUCTIVE TEACHING METHODS EVALUATION, INTEREST-MOTIVATION-SEX ROLE WITH RESPECT TO ATTITUDES, SENSORY AND PERCEPTUAL FACTORS, APPROACHES TO PROBLEM-SOLVING, PERSONALITY VARIABLES, AND MENTAL ABILITIES, AND EXPERIMENTAL ANALYSIS OF BEHAVIOR. THE SECTION OF PROPOSALS PROVIDED A COLLECTION OF PRACTICAL RESEARCH DESIGNS FOR MATHEMATICS TRAINING TO BE PERFORMED WITH LIMITED RESOURCES AND WITHIN A LIMITED AMOUNT OF TIME. THE THIRD AND FINAL SECTION CONTAINED A SURVEY SYNOPSIS OF PSYCHOLOGICAL AND EDUCATIONAL LITERATURE CENTERING ON MATHEMATICS TRAINING DURING THE 1948-58 TIME PERIOD. THAT PART OF THE DATA WHICH WAS NOT GATHERED AS PART OF THE SURVEY WAS OBTAINED FROM PAPERS PRESENTED AT A 2-DAY CONFERENCE ON PROBLEMS AND METHODS IN MATHEMATICS TRAINING HELD AT WASHINGTON UNIVERSITY IN 1959. (JH)

ED002915

JUL 30 1959

OFFICIAL RESEARCH FILE COPY

642

PSYCHOLOGICAL PROBLEMS AND RESEARCH METHODS IN MATHEMATICS TRAINING

Edited by

Rosalind L. Feierabend

and

Philip H. DuBois

Prepared at Washington University, St. Louis,
under Cooperative Research Project No. 642, of
the Office of Education, U. S. Department of
Health, Education and Welfare.

June, 1959

ED002915

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

PSYCHOLOGICAL PROBLEMS AND RESEARCH METHODS IN MATHEMATICS TRAINING

Edited by

Rosalind L. Feierabend

and

Philip H. DuBois

TABLE OF CONTENTS

3414a

Page

INTRODUCTION 1

SECTION I: PROBLEM AREAS IN THE TEACHING OF MATHEMATICS . . . 3

A Suggested Approach to a Systematic Analysis of Teaching-Learning Operations.
C. R. Carpenter 4

Implications of Some Doctrines of Mathematics Teaching for Research in Human Learning.
Robert M. Gagne 11

The Relation of Certain Non-Intellectual Variables to Achievement in Problem-Solving.
Donald W. Taylor 18

The Discovery Method in Training in Arithmetic: A Demonstration.
David A. Page 23

The Role of Nonverbal Awareness in the Learning of Mathematics - Abstract.
Gertrude Hendrix 25

Teaching the Strategy of Problem Solving.
Paul C. Rosenbloom 30

A Rationale for Direct Observation of Verbal Behaviors in the Mathematics Classroom.
E. Muriel J. Wright 38

A Method of Determining the Correlates of Change in Skill in Mathematics.
Philip H. DuBois 47

Scaling Methods Appropriate for Developing Measures of Achievement for Research on Mathematics Training.
Goldine C. Gleser 51

Mental Abilities and Personality Variables in Relation to Proficiency in Mathematics, with Particular reference to the College Level.
S. B. Sells 57

Page

The Interaction of Personality Variables with Teaching Methods Affecting on College Learning.
W. J. McKeachie 62

Perceptual Factors in the Learning of Mathematics.
James M. Vanderplas 66

An Early Approximation to Principles of Programming Continuous Discourse, Self-Instructional Materials.
Thomas F. Gilbert 76

Shortcuts and Analogue Devices in Attempts to Teach the Mathematical Transform Concepts of Feed-Back Theory to Graduate Students in Psychology.
George L. Kreezer 108

SECTION II: PROPOSALS FOR RESEARCH ON THE TEACHING OF MATHEMATICS 116

The Relation of Abstractions to the Concrete: Should Mathematics be Taught in the Context of the Student's Major Area of Interest?
Marion E. Bunch 117

Achievement Determination of Rate of Progression in Learning Mathematics.
C. Ray Carpenter 120

Systematic Exploration of Intra-Personal Correlates of Gain in Mathematics.
Philip H. DuBois 124

Investigation of Differences Between Verbal and Mathematical Conceptualization.
Rosalind L. Feierabend 127

Ability Differences in the Learning of Concepts Governing Directed Numbers.
Robert M. Gagne 131

The Use of an Operant-Conditioning Laboratory for the Investigation of Principles of Programming Self-Instructional Materials.
Thomas F. Gilbert 133

Placement to Maximize Efficiency of Two Methods of Teaching Algebra.
Goldine C. Gleser 136

Page

The Effect of Verbalization of Discoveries on Transfer.
Gertrude Hendrix 138

Empirical Methods for Organizing Mathematics Curricula.
Winton H. Manning 140

The Interaction Between Student Anxiety and Teaching Approach in Learning Mathematics.
Wilbert J. McKeachie 143

Research Problems in Mathematics Teaching.
Paul C. Rosenbloom 145

Some Hypotheses for Psychological Research on Mathematics Training.
Saul B. Sells 154

An Investigation of the Possibility of Improving Attitudes and Achievement of Students by Means of Procedures Designed to Improve Attitudes of Teachers.
Donald W. Taylor 157

Experimental Evaluation of Sensory-Perceptual Aids in Learning Mathematical Concepts.
James M. Vanderplas 160

Patterns in Verbal Behaviors Observed in the Mathematics Classroom and in Individual Pupil Attack of Mathematics Problems.
E. Muriel J. Wright 166

SECTION III: REVIEW OF RESEARCH ON MATHEMATICAL EDUCATION 169

A Review of Research on Mathematical Education 1948-1958.
Rosalind L. Feierabend 170

Bibliography 197

CONFERENCE PROGRAM: Psychological Problems and Research Methods in Mathematics Training 211

PARTICIPANTS 213

INTRODUCTION

A purpose of this report is to provide a handbook or guide for scientific investigators in the field of mathematics training. It is becoming increasingly clear that training in mathematics is a key to the continuance of scientific and technological progress. As a result, various projects for the improvement of teaching mathematics have been established throughout the country, including the School Mathematics Study Group at Yale University, the University of Illinois Committee on School Mathematics, the University of Illinois Arithmetic Project, the Minnesota School Mathematics Center, and the University of Maryland Mathematics Project. These have concentrated mainly on the rewriting of textbook materials and on the use of new principles of teaching.

The contributions which psychology can make to the improvement of mathematics teaching have not as yet been fully assessed nor exploited. It has been the purpose of the present three-month project, sponsored by Washington University in cooperation with the Office of Education, U. S. Department of Health, Education and Welfare under Cooperative Research Project No. 642, to produce an evaluative study of psychological research on the teaching of mathematics.

The report has three main objectives:

1) to formulate problems in the learning of mathematics which might be answered through appropriate experimental investigations;

2) to suggest research methods and designs appropriate to the conduct of such studies; and

3) to review psychological and educational literature, reporting research relevant to the learning of mathematics at all academic levels.

The three sections of this report are concerned with these objectives. The first section, Problem Areas in the Teaching of Mathematics, discusses problems in the field of human learning with special reference to the learning of mathematics. These areas include: problems of measurement, including scaling and the development of appropriate measures of change; the relative value of inductive and deductive methods in teaching; the role of interest, motivation and attitude, and sex differences in respect to attitudes; sensory and perceptual factors; approaches to problem-solving; personality variables and mental abilities in relation to teaching methods; and the experimental analysis of behavior in relation to teaching.

The second section, Proposals for Research on the Teaching of Mathematics, contains a collection of research designs which might be carried out in the field of mathematics training. These designs are intended to be of a practical scope, such that they could be performed with limited resources and within a limited time period.

The third and final section, Review of Research on Mathematical Education, contains a survey of the psychological and educational literature of the past ten years, with special reference to research applicable to the field of mathematics training. Following the review is a list of selected references in this area.

The main event in the course of this project was a Conference on Psychological Problems and Research Methods in Mathematics Training held at Washington University on May 1 and 2. The program and list of participants for the conference are given at the end of this report. Sections I and II of the report contain the papers and research designs presented at this conference. We are extremely grateful to the participants at the conference who gave so generously of their time and thought. By bringing mathematicians and psychologists together for two days to focus on common topics, the conference perhaps awakened a new awareness of each other's approaches to these problems, even though it may not have provided ready answers.

The staff for this project was composed of Philip H. DuBois, Project Director; Rosalind L. Feierabend, Principal Investigator; Marion E. Bunch; Winton H. Manning; and E. Muriel J. Wright, Staff Members; and Portia W. Jones, Secretary.

We would like to extend special thanks to Dr. Randall M. Whaley, Executive Director, Advisory Board on Education, National Academy of Science, for his encouragement and support in initiating and carrying through this project. We are also indebted to Dr. E. G. Begle, School Mathematics Study Group, Yale University, for enabling us to begin preparatory work before the actual date on which the project was inaugurated.

We also owe a debt of thanks to the five assistants who carried out the library research, reading and abstracting the articles for the ten-year period. These were Elizabeth T. Ferrell, Eugene M. Malecki, Carol M. Ridker, Frederick W. Wilke, and Barbara Yaekel.

Our final debt of gratitude is to the two typists of this manuscript, Mrs. Florence Layton and Mrs. E. R. Dunn.

Rosalind L. Feierabend

Philip H. DuBois

Washington University
St. Louis, Missouri
June, 1959

SECTION I: PROBLEM AREAS IN THE TEACHING OF MATHEMATICS

The present section contains the papers which were presented at the Conference on Psychological Problems and Research Methods in Mathematics Training held at Washington University on May 1 and 2. The papers are directed towards problems in the field of human learning, with a special reference to the learning of mathematics. Among the areas discussed are problems of measurement; the use of the discovery method in teaching; the role of non-verbal awareness; the strategy of problem-solving; the relation of mental abilities and personality variables to teaching methods and to proficiency in mathematics; the question of interest in and attitudes toward problem-solving and the existence of sex differences in these attitudes; perceptual factors in learning mathematics; and the implications of an experimental analysis of behavior for research in teaching.

In keeping with the focus of the present report, these papers are oriented towards research possibilities and problems. They are directed to the questions which remain unanswered and which need to be brought under more concerted attack.

The papers are reproduced in the order in which they were presented at the conference.

A SUGGESTED APPROACH TO A SYSTEMATIC ANALYSIS OF TEACHER-LEARNING OPERATIONS

C. R. Carpenter
The Pennsylvania State University

I. Introduction.

Let us begin by confronting a cold, hard and disappointing fact: Predominantly and with distressing frequency, research on single variables of methods of instruction and on conditions of learning has yielded "no statistically significant differences." Are we to infer that differences in methods produce no differences in learning? Is it possible that we have not defined and experimented with those determinants of learning which are most salient and of central importance? Is it possible that we have not arranged for the experimental variables to operate long enough over time to produce significant effects? Or, is it that the determinants of complex learning in academic contexts are so numerous, complex and interactive that effects of single or even several variables are eclipsed by the others in the patterns of relevant determinants? Consideration of this last question would seem to be pertinent as we begin our discussions of psychological problems and research methods in teaching mathematics.

In order to clear the decks still further for vigorous discussion, it is necessary to challenge the semantics of two classes of thought and language about instruction and learning.

First, the ubiquitous practice of formulating educational objectives in highly general, global and abstract terms appears futile, confusing and wasteful, and at times even lacking in integrity. If general objectives must be stated, this should be done with the clear task remaining to be attacked of re-stating the generalizations in more specific terms progressively through n levels of formulation until concrete operations or activities are specifically defined. It must be insured that these operations or activities have syntax with all levels of formulation and with the prime general objectives. It is possible that much of the language used in stating very general educational objectives may be eliminated in the work of designing experiments on learning mathematics and other subjects. The basic questions may be: What actual changes in behavior, or if you prefer, in intellectual development, should be made to occur? What subject matter should interact with the behavior and how is this to be accomplished? Who and what mediates learning of students? How can the effects be measured, the results interpreted and the conclusions employed either in theory testing and formulation or in educational practice?

Second, for purposes of this discussion, it may be desirable and provocative of thought to view critically "learning principles" as they are currently formulated and at this hour are being taught to thousands of prospective teachers. For example, what does the "law of effect" mean to a teacher of mathematics? What does the "principle of reinforcement" really mean for instruction and learning? How can these generalizations be translated into teaching and learning activities? As guides and approaches general laws and principles when integrated into systems of concepts may be useful, but even so, they do not provide forthwith the necessary kinds of intelligible blue prints for education that are so urgently needed if the teaching and learning of mathematics is to be extended, accelerated and improved.

Why am I, at least for the moment, rather disenchanted with experiments on single variables, the pseudo-political activities of formulating very general educational and course objectives and with generalized principles of learning? There are many reasons and rationalizations which could be stated. This, I do not have time for this morning.

Recently I have attended two other conferences; one on the subject of Instrumentation for Learning at Stephens College and the Annual Convention of the Institute of Radio Engineers in New York. An observation may be worth reporting, but regardless of its worth, these recent experiences have triggered one of my current concerns. Educators, psychologists, and design engineers are not talking languages which are mutually understood. The educators and psychologists are not using language or expressing concepts in ways which are understood and which can be translated into electronic or instrument systems by the engineers. The communication engineers in turn have a language that is not understood by educators and psychologists. For me, as a psychologist, the problem is clear but difficult. It is that of how to talk so as to be understood by educators, and thus, to make it possible for educators to translate the language into instructional activities. It is that of how to talk so as to be understood by architects and design engineers, and thus, to make it possible for them to translate the conceptions into instrument systems for mediating instruction for learners.

2. Resolutions of the Problems.

I have dealt at some length with a cluster of problems, the resolutions of which would seem to have considerable professional and perhaps scientific importance. Likewise, the resolutions may have significance for research methods in mathematics instruction.

The preliminary approach which is being suggested involves the very difficult task of systematically conceptualizing all essential operations or functions involved in teaching and learning. To do this, it is required that the specific operations or functions be defined. These may be thought of as constituting sets of operations that can be classified into five or more categories: characteristics of subject matter or content, teaching or instructional functions, mediating or communicating functions, learning functions, and perhaps functions of the environments in which the behaviors occur.

It is suggested that when the essential sets or categories of functions or operations have been defined that specifiable patterns in the different categories can be organized into sub-systems and systems of functions, operations and activities, all of which will have focus on those behavioral changes of students which we call learning.

When this suggested approach is developed and applied in formal education, it may be observed that there is a vast system or complex matrix of determinants, many of which operate in patterns, configurations, sets or facets to influence learning. Furthermore, when the task is completed, it should be possible to locate and place responsibility for the execution of operations and patterns of these. The orders and sequences of events may become intelligible as these eventuate in the academic development of students. It may become clear that single variable research, or even limited patterns of variables, have low probability of producing significant differences in the intellectual behavior or achievements of students. Accordingly, the requirements for designing and conducting research and experimentation in academic learning may be specified with greater realism and accuracy.

Obviously, the problems of quantification and measurement are most difficult. New mathematical and statistical methods will be needed. Great emphasis will be needed on the development and use of research designs which are appropriate to the realities and complexities of the problems. New and expanded research resources and facilities will be needed. Difficult decisions will need to be made between conducting research in context where valid reward systems operate or in the laboratory where experiments can be elegantly controlled.

We of the Division of Academic Research and Services at Penn State have taken some first steps to develop the suggested approach. Our efforts, thus far, have consisted in blocking our three sets of functions under the headings of teaching (t), mediating or communicating (c) and learning (l) functions. We have not yet dealt with context or learning environment functions. Therefore, we are not yet prepared to attack the difficult task of integrating sets of functions into coherent operational systems. We cordially invite your cooperation and constructive criticisms of these preliminary efforts.

Instructional Operations of Teachers (t)

The following is a preliminary list of statements that may be considered descriptive of essential teaching operations. They range over the full scope of an educational enterprise. All of the operations may not be performed in any single classroom, laboratory or teaching situation:

t-1. Developing personal competencies, skills and mastering the resources of subject matter.

t-2. Collecting information, preparing materials, organizing and filing (storing) for availability.

t-3. Instructing students in finding and using resources; materials and favorable conditions for learning.

t-4. Establishing achievement expectancies, learning goals or standards for students and obtaining and providing knowledge of their attainments.

t-5. Interesting, encouraging, involving students and establishing favorable attitudes for intensive sustained effort to learn.

t-6. Programming patterns of information presentations and scheduling related learning behavior of students.

t-7. Gauging the optimum rate of shift of responsibility for learning from the teacher to the student in order to take advantage of his development and to encourage intellectual maturity and independent study.

t-8. Presenting information in varied but appropriate forms to extend its significance and to increase the quantity and quality of learning and retention.

t-9. Providing perspectives for information and responses to be learned, e.g., historical development, theoretical concepts, practical applications.

t-10. Adjusting instruction to students' differences of motivation, interests, previous learning, rates of comprehension and changing abilities to make correct responses.

t-11. Arranging conditions calculated to require appropriate responses of students, e.g., setting problems, stating issues, indicating alternatives for decisions and challenging students.

t-12. Repeating information, with variations, to develop principles, increase generalization, integration and retention of information and skills.

t-13. Serving as an authoritative commentator, critic and interpreter of the subject and of the learning behavior of the student.

t-14. Counselling and guiding students in and through a subject, course of study or curriculum.

t-15. Evaluating, correcting, confirming or extinguishing, with optimum psychological timing, the trial performances of students.

t-16. Assessing the achievements of students relative to expectancies, potentialities, objectives and standards.

t-17. Explaining the personal relevance of information to students and learning to know the unique intellectual characteristics of individuals.

t-18. Serving as a model for students as a person, professional man, and scientist or scholar.

Mediating Communications Network Operations (c)

Continuing the exploratory preliminary approach undertaken here, the next step is an analysis and brief description of possible operations of communication systems for mediating instruction and learning. The following are suggested as an initial list:

c-1. Information collecting operations.

c-2. Information or instruction storage (filing) operations.

c-3. Information sorting, selecting and identification.

c-4. Information ordering, organizing and recording operations.

c-5. Regulating forms, amounts and rates of presentations and regulating appropriate responses.

c-6. Transmission and routing of information over channels and distances to learners and users.

c-7. Information presentations of various types and modes.

c-8. Regulating, focusing and shaping complex materials to aid perception and responses.

c-9. Implementing intrinsic and extrinsic reward systems for reinforcing learning.

c-10. Extending, expanding and reducing learning cues, signs and symbols.

c-11. Translating, varying and repeating instruction.

c-12. Relating abstract or theoretical learning to reality and applied situations.

c-13. Recording of responses and effects, and displaying the records to those who need them.

c-14. Examining and measuring response capabilities of students and validating these against sets of performance criteria.

c-15. Recording, reporting, analyzing and summarizing the results of learning.

The Operations and Conditions of Learning of Students (L)

The third domain of operations, in our hypothetical systems approach, is that of student actions in learning situations. Here there are two basic questions: First, what are the essential operations and conditions necessary for learning to occur? And, second, beyond the basic essentials, what are the operations and conditions which tend toward maximizing learning for a given student of a given subject? The list that follows is again exploratory and is intended to illustrate another possible aspect of the general problem:

L-1. Selectively activating the relevant and previously learned information, skills and responses.

L-2. Discovering sources of materials and learning opportunities.

L-3. Establishing levels of expectancies for learning and achievement in terms of response possibilities.

L-4. Mobilizing energy, directing and sustaining adequate levels of effort for learning.

L-5. Scheduling learning activities and ordering content for effective learning.

L-6. Establishing and maintaining necessary perceptual contact with subject content.

L-7. Defining and limiting the learning demands in terms of information to be learned, and the knowledge and responses desired or required.

L-8. Developing sensitivity to extrinsic and intrinsic reward systems for reinforcing achievement.

L-9. Acquiring the instrumental skills essential for the learning of information and responses of different kinds and levels of complexity.

L-10. Applying learned responses appropriately to meet performance requirements.

L-11. Reviewing, repeating and confirming new content and responses.

L-12. Practicing the making of discriminations, judgments and decisions.

L-13. Practicing the selective elimination of the distracting and irrelevant in learning situations.

L-14. Integrating and organizing elements of knowledge and skills into functional systems of understanding and performances.

L-15. Achieving abilities for symbol translation, transformation, condensation, generalization and application.

L-16. Judging indicators of directions and rates of change in response acquisitions.

L-17. Testing, checking and validating gains in information and skills.

L-18. Modeling the behavior and role characteristics of appropriate prototypes of teachers.

IMPLICATIONS OF SOME DOCTRINES OF MATHEMATICS TEACHING FOR RESEARCH IN HUMAN LEARNING*

Robert M. Gagné
Princeton University

One can hardly read any part of the literature on the teaching of mathematics, whether current or of the more recent past, without being struck by the fact that there are certain rather basic points of view which have attained a degree of popularity. Each of them is supported by some evidence, ranging from simple classroom observation to relatively systematic experimental findings. Each has some importance, apparently, in determining the basic assumptions of a smaller or larger group of mathematics educators towards certain kinds of curricular content, or certain kinds of methods of presentation of material to be learned.

To the learning psychologist, these points of view often represent the fundamental issues in human learning. That is, they are the questions about how people learn--about the nature of the theoretical constructs which must be used to explain learning and problem-solving performance. From his point of view, therefore, these ideas represent assumptions about how human behavior is organized, or, if you will, about how the human nervous system works. To the learning psychologist, these ideas raise questions of theory which need to be investigated by means of systematic thinking and research.

It may be worthwhile at the outset to point out that the kind of assumptions I am talking about here seldom arise from, and are just as seldom verified, by practical tests of curricular content in the classroom. When a mathematics educator has an idea about a change in content of mathematical knowledge to be taught--and this too appears to be a popular game nowadays--he nearly always recognizes the need for a practical test of its feasibility and worth. Can number theory be successfully taught in the elementary grades? Can solid geometry be taught as a part of plane geometry, or set theory as a part of algebra? Can number concepts be introduced by the use of blocks and other concrete objects, rather than by symbols? There can be little doubt of the usefulness of such practical classroom evaluations of new materials, nor of the desirability of making such tests as objective and powerful as they can be. But it is rare indeed when such tests provide acceptable evidence for or against the basic assumptions about human learning on which the changes may have been initially predicted. The reasons for this are fairly well known. The kinds of variables being manipulated in such tests have a complex nature--they are not the kind which possess directly defined relationships between experimental operations and hypothetical constructs demanded by the formulation of scientific theories of human behavior. To get at these basic assumptions we must have precisely defined operations which specify, on the one hand, how the processes of behavior are modified by manipulation of the situation, and on the other, what kinds of human performance result from such modification. It is primarily in this way that a behavioral experimental study differs from the usual classroom test.

---

*This report was made possible in part by funds granted by Carnegie Corporation of New York. That Corporation is not, however, the author, owner, publisher or proprietor of this publication, and is not to be understood as approving by virtue of its grant any of the statements made or views expressed therein.

What I should like to do here is to take two or three examples of important ideas in mathematics teaching, and to define, if I can, the sorts of questions these raise about the nature of human learning and performance. I shall try to show that these questions are ones about which we do not know the answers. They are assumptions, not laws. At the same time, I hope to be able to suggest the kinds of experimental studies which could be expected to approach these questions in a serious and systematic way.

The Role of Language

As a first problem, let me consider the role of language in the learning of mathematics. What do words, the English language, the "verbalization of rules" have to do with mathematics learning? There are many who think that words interfere with learning. For a summary of these ideas, I refer to a chapter on "Transfer of Training" by Rosskopf (1953) in the Twenty-First Yearbook of the National Council of Teachers of Mathematics, as well as to the ideas expressed in the textbook in High School Mathematics of the University of Illinois Committee on School Mathematics (1957). As a concrete example, take the verbal rule "Like signs give plus" (Brune, 1953). What is wrong with such a verbalization? Is it bad for problem solving because it is expressed in words? Or is it bad because it does not have enough words?

A notable example of the idea that verbalizing interferes with learning occurs in the Illinois textbook referred to above. Here, the addition of directed numbers is undertaken by attempting to teach students what might be called a "spatial model": taking a trip. The student is asked to imagine taking a trip in one direction when the sign is plus, and in the other when the sign is minus. The sum of two directed numbers is obtained by finding the distance between the point where he started his trip and the point where he ended up (assuming that the second trip began where the first ended). Thus the sum of +5 and -3 is to be arrived at directly by imagining that he moves in these directions, and thus ends up in the direction +, and 2 units from where he started. There is little doubt that this is both an accurate and effective way of arriving at the sum of two signed numbers.

But why is it a better way? The answer to this is not at all obvious. Is it because the verbal rules are harder to learn—rules like "If the two numbers have opposite signs, subtract and affix the sign of the larger"? There would appear to be much doubt about such a reason. The student already knows the meaning of all the words in such statements—words like "opposite", "sign", "subtract". The committing of such verbal rules to memory is surely a simple task, particularly if we do not demand exact literalness of reproduction. Is it because the acquired spatial model is more accurate? Are there some instances in which the spatial model applies correctly, whereas the verbal rule does not? I cannot see that this is so, as long as we are talking about the summing of directed numbers. Is it because the spatial model leads to a superior "understanding of symbols" as compared to a verbal rule? This is indeed a question which could be subjected to an experimental test, but so far as I know, it has not been. Of course, in doing such an experiment, one would first have to define exactly what is meant by "understanding of symbols".

It is possible to perceive potential disadvantages, even, in the acquisition of a spatial model. Having acquired such a model, will the student be able readily to use it in solving problems about profit and loss, gain and loss in weight, changes in temperature, increases and decreases in time? Is "taking a trip" really an effective mediating process for the solution of a problem involving the combination of positive and negative values of velocity or acceleration, where the possibility of confusion between the "distance" of the model and distance as a physical variable seems potentially great? I do not think we know this. Nor do I want to suggest here that students cannot learn this sort of mediating process—obviously they can and do, at the University of Illinois High School. The question it poses is not whether it can be done, but rather how and why such an acquired mediating process works in human thinking.

Before leaving this important topic, I want to emphasize again that we should not lose sight of the tremendous power language has as a mediator of human performance. We should all agree, I think, that learning in mathematics could not take place without language. Think of how much space is taken up in a mathematics textbook by the English language, or how much is used by the teacher in fostering the learning of new concepts and rules. Many criticisms of the limitations of verbalization in mathematics, it seems to me, raise the question "Are verbal rules said to be ineffective in certain areas of mathematics because they are verbal, or possibly because they use the wrong words"? I do not know the answer to this question, but I think it is one which could be subjected to research of a rather fundamental sort.

Let me distinguish here between two possible meanings of "verbalization". Given the problem of solving for x in the expression

$$2 = \frac{10 + x}{x},$$

it seems to me that two of the most useful concepts which can mediate the behavior required in taking the first step are the words "denominator" and "multiply". These are (presumably) well-learned words, readily available to the new student of algebra, which can control a tremendous variety of appropriate responses. They therefore appear to be extremely useful concepts in this particular step of the problem-solving process. On the other hand, the learning of a formal "rule" in all its logical elegance may indeed be a process of little value to the solution of problems. For example, the principle

$$\frac{1}{x} \; \frac{1}{y} = \frac{1}{xy}$$

may be expressed in verbal terms as "The product of the reciprocals of two numbers is equal to the reciprocal of their product". Certainly one is entitled to doubt the value of this form of "verbalization".

The basic theoretical questions raised by these ideas about language and mathematics may be summarized as follows:

1. How do words function in the learning and use of mathematical principles—as stimuli, as labelling responses, and as concepts?

2. Are there differences in the adequacy of words and visual models as mediators of rules and problem solving?

3. What is the process that accounts for the "dropping out" of verbal mediation in the solution of problems symbolically?

4. If words are used in mediating problem solving, what conditions determine the learning of precise meanings for these words, so that they do not lead to errors in performance?

Basic Abilities in Mathematics Learning

As a second example, I want to discuss the related topic of basic abilities. The existence of individual differences, presumably relatively stable ones of the sort implied by the term "basic abilities", is a pervasive issue in mathematics teaching, as well as in mathematics itself. It is a widely held belief, for example, that some people tend to solve mathematical problems by thinking out the solutions in terms of visual and spatial models and then translating these into acceptable symbols, while other people think out solutions in terms of abstract symbols and symbol-operations. It is thought possible by some that problems may be solved, by some people, in terms of logico-verbal statements, and then translated into symbolic form. With regard to these differences, it is probably significant that some of the best-defined and stable differences in basic abilities which have been measured by psychologists are spatial factors (or ability), verbal factors, and numerical factors.

What do such abilities have to do with mathematics learning? One implication that may readily be appreciated is that a measure of "numerical ability" should predict the relative success of students in mathematics courses. This it certainly does, to a relatively high degree. But the basic logic of ability measurement should enable us to do much more than that. Basic abilities are assumed to have the property of supporting or facilitating the learning of a great variety of specific skills. Thus the reason for believing in the existence of a specific "spatial ability" is actually that it has been shown to be correlated with performance on a large number of different specific spatial tasks, each of which requires the learning of new and different behavioral processes.

If this is so, then one should be able to verify a number of hypotheses, not only about the nature of abilities involved in mathematics, but also about the essential nature of what is learned in mathematics. The kinds of hypotheses I mean are the following:

1. People who are high in "spatial ability" should acquire mediating spatial concepts more readily than they do symbolic or verbal ones. Similarly, for "numerical ability" and "verbal ability".

2. People who are high in "verbal ability" should reveal a better performance on geometric problems when they are taught verbal mediators than when they are taught spatial or symbolic ones. Similar statements could be made for the other basic abilities, "spatial" and "numerical", using algebraic and formal logic problems as performance criteria.

3. When people high in "spatial ability" are taught the same mathematical principle in terms of spatial, verbal, and symbolic concepts, measuring retention of this principle after a period of time should reveal increasing dependence on spatial concepts. Again, similar hypotheses could be made to pertain to the other abilities.

Experimental investigation of problems like these holds out the promise of making aptitude measurement a more powerful technique than it is at present, not only for practical purposes, but in terms of improved understanding of human behavior. The crucial approach here is one of devising basic hypotheses, not simply on the prediction of correlations between two types of performance, but rather between an inferred basic ability and the improvement of some specific conceptual behavior. If the fundamental logic of ability measurement is sound, we should be able to make predictions not simply about total performance in a subject like mathematics, but also about the specifically learned concepts which mediate this performance.

Discovery

As a third example, I should like to consider the very important notion of discovery and its role as a method of learning in mathematics. This idea, having its origin partly in the writings of members of the Gestalt school of psychology (Katona, 1940, Wertheimer, 1945) has had a profound influence on mathematicians and mathematical educators. Basically, the notion is that productive thinking, or inventive problem solving, is achieved by an "insight" which represents an essential reorganization of the entire field in which the individual and the problem are located. What brings about this reorganization is a set of factors which tend to structure the insightful process—like closure, pragnanz, symmetry, and others. It is becoming pretty clear nowadays that these ideas, whatever has been their usefulness in the stimulation of thought, have led neither to a consistent theory of behavior nor to a body of psychological knowledge which can be used to back them up.

I know of no psychologist, behaviorist or otherwise, who ever denied the existence of insight as a fact, describing the sudden change in behavior which occurs when the varied activities of "wrestling with a problem" give way to the "achievement of solution". Certainly there is a sudden change in behavior, and certainly this must represent a "reorganization" of underlying behavioral processes. If there is any issue here to be argued about, of course it must be the question, What are the underlying behavioral processes?

Nor has there been any argument (unless it be in the imagination of certain writers) about the aims of mathematics training. Is a course in arithmetic given so that the student can solve all the problems in the textbook, or is it given so that he will be able to solve new problems which he has never encountered before? Of course it is the latter. Everyone agrees, I think, that mathematics is taught with the purpose of making it possible for the student to apply the concepts and rules he acquires to novel situations, to "invent" solutions to newly encountered problems, whether these be in physics, in some more advanced mathematical subject, or even in his everyday life. I should judge that, if there is a legitimate question to be answered here (that is, a question of interest to the psychologist) it is, What must the individual learn in order to give him the capability of solving novel problems? In other words, what kind of learning will make the student capable of discovery?

There is one answer to this problem which the psychologist, while maintaining his objectivity, must nevertheless be highly suspicious of, it seems to me. This may be stated as follows: the way to make the student capable of discovery is to have him practice discovery. (The way to teach the capability of solving new problems in geometry, for example, is to continuously confront the student with new problems, or "originals".) The reasons for one psychologist's doubts about this principle are two:

1. Carried to a logical extreme, it is obviously false. An individual cannot solve a truly original problem without having acquired a considerable background of concepts and rules. If he doesn't know about similar triangles he cannot supply a geometrical proof requiring this set of principles, no matter how hard or long he tries. Thus "originals" are not really new problems, but only new in a certain sense. If we could define in what sense they are new, we might get closer to the heart of the problem. Meanwhile, the effects of negative reinforcement on those students who do not manage to "discover" when they are asked to "discover" are not considered likely to be beneficial for learning.

2. There are many instances of learning other types of behavior in which the principle "the best way to learn is to practice the ultimate task" has been shown to be false. For example, the best way to learn a motor performance is not always to practice the task in its final form. "Discovery" of the correct stimuli can often be facilitated by practice which does not have the form of the final performance. Industrial workers may learn an operation like disc-cutting (Lindahl, 1945) by first practicing patterns of movement. Pilots may learn the performances required in flying by practicing underlying skills in a training device. Or take another example from the conceptual field. Thoughtful students of the subject of reading are now pretty well agreed that the best way to teach Johnny to read, particularly in the sense of reading novel words, is to teach him to pronounce letter combinations (Brown, 1958). To a large extent, what the child learns (concepts of pronunciation) is not the same as what he does (read new words). If we want a student to be able to discover answers to new mathematical problems, is the best method of learning "practice in discovery"? I do not know, but you will understand my doubts.

The kinds of basic questions suggested by this discussion of the doctrine of discovery, then, are these:

1. What are the dimensions of novelty in problems? What makes a problem new (so that the answer must be discovered) as opposed to familiar (so that solution is simply a matter of applying a known rule)?

2. To what extent can success in problem solving (novel or otherwise) be related causally to the adequacy of mastery of mediating concepts? To what extent must "field organization" factors be used to explain problem solving?

3. What are the factors which affect the degree of generalizability of concepts, and how are these related to problem-solving behavior?

Now, to review what I have tried to say in this paper. First, there are certain prevalent ideas in mathematics teaching which raise basic issues of behavior theory. The answers to these questions can best be sought by means of theoretically based experimental studies. Three of the most important ideas are those of the role of language in mathematics learning, the effects of individual differences in basic abilities, and the doctrine of discovery. In the case of these notions, I have tried to show that they are really basic questions about behavioral organization to which we have as yet no dependable answers. Each of them raises a number of problems which could be attacked by the experimental psychologist, and which, if solved, would undoubtedly be significant contributions to our knowledge of how people learn to solve problems with the use of mathematics.

REFERENCES

Brown, R., Words and things. Glencoe, Ill.: The Free Press, 1958.

Brune, I. H., Language in mathematics. Ch. 5 in The learning of mathematics: its theory and practice. Twenty-First Yearbook. Washington, D.C.: National Council of Teachers of Mathematics, 1953.

Katona, G., Organizing and memorizing. New York: Columbia Univ. Press, 1940.

Lindahl, L. G., Movement analysis as an industrial training method. J. Appl. Psychol., 1945, 29, 420-436.

Rosskopf, M. F., Transfer of training, Ch. 7 in The learning of mathematics: its theory and practice. Twenty-First Yearbook. Washington, D.C.: National Council of Teachers of Mathematics, 1953.

University of Illinois Committee on Mathematics, High School mathematics. First course. Teachers edition. Urbana, Ill.: University of Illinois Committee on School Mathematics, 1957.

Wertheimer, M., Productive thinking. New York: Harper, 1945.

THE RELATION OF CERTAIN NON-INTELLECTUAL VARIABLES TO ACHIEVEMENT IN PROBLEM SOLVING

Donald W. Taylor
Yale University

During the past eight years, first at Stanford and more recently at Yale, my research assistants and I have been engaged in a program of research on thinking, with support provided by the Psychological Science Division of the Office of Naval Research. As one major part of that work, more than a dozen studies have been carried out to explore the relation of certain non-intellectual variables to achievement in problem solving. I should like to review with you the results of several of those studies and then examine what I believe to be their implications for the teaching of mathematics.

This exploration of the role of non-intellectual variables actually began with a series of four studies by Sweeney (1953) concerned with sex differences in problem solving. Previously, extensive research had been devoted to an examination of possible sex differences on tests of intelligence. A careful review of such research had led Terman, et al. (1946) to conclude: "No positive claim can be made for the superiority in general intelligence of either sex at any age." This conclusion is, I believe, rather generally accepted among psychologists. Nevertheless, an examination of the research literature prior to Sweeney's work reveals several instances in which significant differences between men and women were obtained in achievement in problem solving, obtained usually incidentally in the course of experiments designed to test other hypotheses.

Sweeney's studies (1953) taken together support the conclusion that for at least certain kinds of problems the mean performance of men is significantly better than that of women, and that this sex difference remains significant when the groups are matched with respect to intelligence, various special abilities, or relevant specific knowledge. Thus, for example, in one study he employed a simple and a complex version of a single problem; the complex differed from the simple version in that it involved an additional restriction which necessitated an additional detour step for solution. Groups of men and women were trained to the same criterion of performance on the simple version and hence presumably to the same level of relevant knowledge. When subsequently tested on the complex version, men performed significantly better than women. In a related earlier study, Billings (1934) had given groups of men and women the necessary information to solve problems in mechanics, economics, geometry and history, and later tested for retention of the material. The men and women did not differ in amount of information retained or in tested intelligence, but when they solved problems in which this information was to be utilized, the men's mean score was considerably higher than that of the women. In a more recent study, Kostick (1954) also obtained data supporting the conclusion that the sex difference in achievement cannot be accounted for in terms of available relevant knowledge.

In his most elaborate study, Sweeney (1953) employed matching to control differences not only in intelligence as measured by the second part of the Terman Concept Mastery Test, but also other variables upon which the sexes are known to differ, including visualization ability, mechanical comprehension, mathematics background, and mathematical achievement. A variety of problems were given the subjects, including tool problems, puzzle book problems, and mathematics-logic problems. On some of the problems, matching on intelligence or on the other variables resulted in disappearances of sex differences. But on other kinds the differences favoring the men remained highly significant even after such matchings. In general, sex differences disappeared for matched groups when the problems involved were relatively straightforward, but remained significant when the problems required restructuring or reorganization for solution. Subsequent work by Carey (1955), Nakamura (1955), Milton (1957, 1958), and Berry (1958) has yielded significant differences favoring men for groups of men and women which did not differ in intelligence or in which the influence of such differences were controlled by the use of partial correlation or of analysis of covariance.

Accounting for this sex difference in problem solving is of interest in itself, but it is of even more interest in that whatever variables operate to produce this difference may also be presumed to influence achievement in problem solving among members of either sex.

Carey (1955) obtained data indicating that the inferior achievement of women may result from the fact that they have less favorable attitudes. Employing Likert-scaling procedures, she constructed two forms of a scale designed to measure attitude toward analytic thinking. When tested with this scale, Stanford undergraduate men scored significantly higher than did Stanford undergraduate women; such scores correlated with achievement in problem solving among men. In subsequent studies by Milton (1957) and by Berry (1958, 1959) significant correlations have been obtained between scores on this same attitude scale and problem solving, and in several cases the correlation remained significant when the influence of mathematical and verbal aptitude was removed by partialling. But these results do not indicate which came first—the chicken or the egg—the more favorable attitude or the superior achievement.

To test the influence of attitude upon performance more directly, Carey (1955) undertook to change attitude experimentally. Groups containing equal numbers of men and women participated in unstructured discussions in which they were asked for help in suggesting what factors might influence achievement in problem solving. These sessions involved primarily discussion, usually initiated by some member of the group, of the role of attitude or interest in problem solving and exploration of the attitudes of individual members toward problem solving. No mention of sex differences was ever made. After such discussions, the subjects were given a second form of the attitude scale and a second measure of achievement in problem solving. The interesting finding was that, as predicted, the achievement of the women increased significantly more than that of the men; this she predicted because she believed that the initially poor attitude of the women would be more subject to improvement than the initially favorable attitude of the men. It must be noted that though the predicted increase in performance occurred, no significant change in the attitude scores was obtained for men or for women. The *post hoc* explanation is that the types of items employed on the attitude scale were relatively insensitive to short-term change though quite adequate as a measure of general level of attitude.

The relation to problem solving of other non-intellectual variables has also been investigated. Employing the apparatus devised by Crutchfield (1955) for measuring the tendency of individuals to conform to group opinions, Nakamura (1955) measured the tendency to conform of a large number of undergraduate men and women at the University of California. As he expected, the women showed significantly more tendency to conform than men, a difference which he thought might be related to the sex difference in problem solving. He also administered to the same subjects a measure of achievement in problem solving and a test of intelligence. Of most interest is his finding of a significant negative correlation among men between problem solving and conformity after the influence of variation in intelligence had been partialled out. This same correlation for women was also negative, but not significantly different from zero. The sex difference in problem solving remained significant when an analysis of covariance was carried out with intelligence and conformity as control variables.

As a part of a much larger study, Berry (1959) obtained for a group of Yale undergraduate men measures of tendency to conform, achievement in problem solving, and of mathematical and verbal aptitude. When he computed the partial correlation between conformity and problem solving with mathematical and verbal aptitude as control variables, it did not differ significantly from zero. He thus failed to confirm Nakamura's principal finding. More work will be necessary to clarify any possible relation between tendency to conform and achievement in problem solving.

The results obtained by Sweeney (1953), Carey (1955), and Nakamura (1955) led Milton (1957) to suggest that the sex difference in problem solving may result primarily from differences between the roles appropriate to men and to women, and further, that within either sex achievement in problem solving will be significantly correlated with degree of identification with the masculine role. In his first study (Milton, 1957) employing the Terman-Miles Attitude-Interest Test (and also the M-F score on the Minnesota Multiphasic Personality Inventory as well as a Behavioral Inventory) as a measure of sex-role identification, he obtained significant relations between masculinity and problem solving both among men and among women, even after variation associated with mathematical and with verbal aptitude had been appropriately controlled. The subjects employed were Stanford undergraduate women. In two more recent studies, Berry (1958, 1959) failed to confirm this finding of Milton's. The first of these involved a large group of Swarthmore undergraduate men and women. The second was the one involving Yale men already referred to above. In both studies, scores were obtained for verbal and mathematical aptitude, achievement in problem solving, and for the Terman-Miles Attitude-Interest Test. The correlation between degree of masculinity and problem solving with the effects of mathematical and verbal aptitude partialled out failed to reach significance for the Yale men or for either the Swarthmore men or women. This discrepancy in findings cannot be fully explained. It should be noted that in another study which employed high school students as subjects, Milton (1958) did find that among boys M-F scores added significantly to the prediction of problem solving by a test of intelligence, although this was not true among girls.

Milton (1958) also carried out two studies in which he explored the influence of the content of problems upon the sex differences in problem solving. One study employed as subjects undergraduate men and women at

Stanford University and the other undergraduate men and women at the University of Colorado. Twenty problems were given in each study, every problem being presented both with masculine content and with feminine content. Any given subject received half the problems in masculine and half in feminine form. As predicted, presenting the problems in feminine form resulted in a significant reduction in the sex differences. It should be noted, however, that in both cases a small difference remained favoring the men, this difference being significant in the second but not the first study.

More work will certainly be needed to clarify the role of non-intellectual variables in problem solving, but the work completed thus far has, I believe, certain implications worth examining for those interested in the teaching of mathematics.

The finding which is most firmly established by the studies described here is that on the average men do better than women in solving at least certain kinds of problems and that this difference cannot be fully accounted for in terms of differences in general intelligence, mathematics aptitude, relevant specific knowledge, or various other specific aptitudes. The populations from which samples were drawn in these studies demonstrating this sex difference include college undergraduates at Stanford University, the University of California, Swarthmore College, and the University of Colorado, and also students from a high school in California and from a high school in Connecticut. Carefully controlled studies of possible sex differences in problem solving are not available at the elementary school level. What data is available suggests that it is doubtful that a sex difference of the same magnitude will be found at least in the early grades. If such a difference is not found in the early grades, what then happens during the later elementary grades and possibly the first years of high school to produce the sizable sex difference found in the studies described—a sex difference in solving not only problems with mathematical content, but also spatial and verbal reasoning problems?

The studies summarized also point strongly toward the possible influence of attitude upon achievement in problem solving. Other data (Alpert, personal communication) show that during the school year attitudes of elementary school students tend to shift toward the attitudes of the teacher. Less systematic information would suggest that a sizable majority of elementary school teachers, most of whom are women, have distinctly unfavorable attitudes toward mathematics and toward analytic thinking more generally. If we may assume that girls will be more likely to identify with, and hence to be influenced by, the attitude of the women teachers, we may suppose that the sex differences in problem solving observed at the high school and college level is at least in part the result of less favorable attitudes acquired in the later elementary and early secondary years. (Even should this turn out to be demonstrable, it would not, of course, exclude the possible contribution of other factors to the observed sex difference.) It may also be supposed that the boys may be influenced, though less so, by the teachers with unfavorable attitudes and hence subsequently do less well in mathematics and in problem solving generally than their abilities would enable them to do.

You have received a copy of a proposal which I have made for "An Investigation of the Possibility of Improving Attitudes and Achievement of

Students by Means of Procedures Designed to Improve Attitudes of Teachers."* May I emphasize that should this experiment yield positive results and thus demonstrate the influence of attitudes of teachers upon achievement of students in problem solving, I would not then propose that a program be undertaken designed to improve the attitudes of large numbers of teachers toward mathematics or toward analytic thinking. This, I believe, would be much too difficult an undertaking. I would propose instead that emphasis should be placed upon having mathematics taught throughout the elementary grades by specialist teachers—teachers chosen not only for their competence but also because they are among the minority who already have the desired favorable attitudes.

REFERENCES

Berry, P. C. (1958) *An exploration of the interrelations among some non-intellectual predictors of achievement in problem solving.* New Haven: Depts. Industr. Admin. & Psychol., Yale Univer. (Tech. Rep. 4, Contract Nonr 609-20).

Berry, P. C. (1959) *A second exploration of the interrelations among some non-intellectual predictors of achievement in problem solving.* New Haven: Depts. Industr. Admin. & Psychol., Yale Univer. (Tech. Rep. 4, Contract Nonr 609-20).

Billings, M. L. (1934) Problem solving in different fields of endeavor, *Amer. J. Psychol.*, *46*, 259-272.

Carey, Gloria L. (1955) *Reduction of sex differences in problem solving by improvement of attitude through group discussion.* Stanford: Dept. Psychol., Stanford Univer. (Tech. Rep. 9, Contract Nononr 25125). Also appeared as: Carey, Gloria L. (1958) Sex differences in problem solving as a function of attitude differences. *J. abnorm. soc. Psychol.*, *56*, 256-260.

Crutchfield, R. S. (1955) Conformity and character. *Amer. Psychologist*, *10*, 191-198.

Kostick, M. M. (1954) A study of transfer: sex differences in the reasoning process. *J. educ. Psychol.*, *45*, 449-458.

Milton, G. A. (1957) The effects of sex-role identification upon problem solving skill, *J. abnorm. soc. Psychol.*, *55*, 219-244.

Milton, G. A. (1958) *Five studies of the relation between sex-role identification and achievement in problem solving.* New Haven: Depts. Industr. Admin. & Psychol., Yale Univer. (Tech. Rep. 3, Contract Nonr 609-20).

Nakamura, C.Y. (1955) *The relation between conformity and problem solving.* Stanford: Dept. Psychol., Stanford Univer. (Tech. Rep. 11, Contract N6onr 25125). Also appeared as: Nakamura, C. Y. (1958) Conformity and problem solving. *J. abnorm. soc. Psychol.*, *56*, 315-320.

Sweeney, E. J. (1953) *Sex differences in problem solving.* Stanford: Dept. Psychol., Stanford Univer. (Tech. Rep. 1, Contract N6onr 25125).

Terman, L. M., Johnson, Winifred, Kuznets, G. & McNemar, Olga W. (1946) Psychological sex differences. In Carmichael, L. *Manual of child psychology.* New York: Wiley.

---

*See the Research Design by Donald W. Taylor published in this report.

THE DISCOVERY METHOD IN TRAINING IN ARITHMETIC: A DEMONSTRATION*

David A. Page
University of Illinois Arithmetic Project

[One of the highlights of the Conference on Psychological Problems and Research Methods was a 45-minute demonstration by Professor Page of the discovery method in teaching arithmetic. The nature of the mathematical problem employed is discussed below by Professor Page. The participants in the demonstration were some fifteen sixth-grade pupils from the Glenridge School in Clayton, Missouri, who were chosen for their superior mental ability.** They were totally unfamiliar, both with the problem and with the demonstrator.

The mathematical question was first presented in terms of a set of blocks with which each child was provided, and the pupils were asked to phrase the problem in their own words and to agree upon their terms. Once some degree of understanding had been reached, the students transferred their work from the blocks to the blackboard. Towards the end of the demonstration, an attempt was made to have the pupils state the mathematical principle in general verbal terms.

R.L.F.]

The demonstration class was given to attempt to show some principles which could not be described adequately otherwise. Consequently this will summarize only certain aspects of the class.

A ten year old 4th grade farm boy in California suggested the problem which was investigated during the class. Stated quickly the problem was this:

> Definition:
> Among the positive whole numbers, call $n + 1$ and $n - 1$ the first neighbors of n. Similarly, call $n + 2$ and $n - 2$ the second neighbors of n, etc.
>
> Question:
> Given the numbers 1, 2, and 3 in order, can you rearrange them so that no number is next to either of its first neighbors?

---

*Prepared in part under a grant from the Carnegie Corporation of New York to the University of Illinois Arithmetic Project.

**We are extremely grateful to Miss Virginia Jackson of the Glenridge School in Clayton, Missouri who kindly agreed to provide the pupils for the demonstration.

Answer:
No. However it can be done with 1, 2, 3, and 4. For example: 3, 1, 4, and 2.

Question:
Using consecutive integers beginning with 1, how many are required in order that they can be rearranged so that each number is kept away from its first neighbors, second neighbors, and so forth up to kth neighbors?

The students discovered that the answer to the second question is 2 k + 2 (in different language).

The topic was chosen not because it is in the main stream of any existing or proposed course but because it is sufficiently self-contained to be dealt with in one period and provide an opportunity for student discovery.

* * *

The University of Illinois Arithmetic Project is preparing new elementary mathematics courses for grades 1 through 6. Before preparing year long courses it is producing a variety of units on special topics which can be used to augment present courses and which will serve as raw material for the ultimate design of new courses. Inquiries are invited.

THE ROLE OF NONVERBAL AWARENESS IN THE LEARNING OF MATHEMATICS—Abstract

Gertrude Hendrix
University of Illinois School Mathematics Project,
University of Illinois

For any up-to-date program in Secondary School Mathematics the retraining of teachers presents a unique problem in that simple but profound and very subtle concepts are being taught to unsophisticated high school students. To attain an operational mastery of these penetrating concepts and generalizations, young students must acquire them through nonverbal awareness, much as the concepts and generalizations must have "dawned upon" the mathematician who first grasped them. Instructional materials can be written to promote this kind of learning, but training teachers so that they will capitalize upon and not misuse the materials requires that the teachers be aware of the phenomena of learning through nonverbal awareness. Learning to teach so that this kind of awareness is developed and learning to recognize when a student has it are two great difficulties confronting teachers of these new mathematics courses.

Occasionally, one finds someone who picks up the knack very quickly. Such a teacher is usually one who has developed his own way of "getting things across to" very small children, or someone with a knack for teaching animals, or (in several outstanding cases) a creative mathematician or scientist who through introspection recognizes the process as a familiar concomitant of his own experience in research. All others must have an opportunity to see and feel this kind of learning going on. If one tries to tell them about it, they think that he is talking about something else—or else talking about nothing.

Even after a teacher has overcome the two difficulties cited above, there remains the great problem of reconciling this new knowledge of how people learn with whatever else the teacher knew (or thought he knew) about learning and instruction. This reconciliation poses many basic problems of great importance. Attached is a list of questions and hypotheses prepared for another conference in April, 1958. The questions are as yet unanswered. Their treatment would require a long-range research project and a research team of the highest caliber—a team that would enlist the best to be found in the fields of anthropology, genetics, ethology, linguistics, communication, and logic, as well as from psychology.

PSYCHOLOGICAL PROBLEMS PERTINENT TO MATHEMATICS EDUCATION

CRITICAL UNANSWERED QUESTIONS

1. Why is something which has come to a person first as insight (that is, non-verbal awareness) much more likely to "pop into his mind" when he needs it than something which he has acquired through interpreting another's linguistic formulation?

2. One of the ways in which a student attains* a concept is by grouping instances of it through association with the same word or symbol.** To what extent (if any) does this manner of attainment handicap him in recognizing instances of the concept

   a) if he does not know that symbols are assigned to their referents instead of belonging to them naturally?

   b) if he does know that relation of symbol to referent is artificial?

3. Compare the probabilities of one's incorporating a newly-learned word (or symbol) into his "production vocabulary"*** if the word or symbol comes to him

   a) as a name for a concept he has already attained to the non-verbal awareness stage, or

   b) as a word or symbol whose referent he acquires from context, or

   c) as a word heard or seen and looked up in a dictionary?

4. a) Under what conditions does attempting to verbalize a concept immediately after one has attained it mutilate the concept in his own mind?

---

*By 'attaining a concept' the writer means acquiring a way of grouping all entities into those which belong (to the group) and those which do not belong. For example, one who has the concept named 'telephone' can classify each object he encounters as a telephone or not a telephone; one who has the concept of odd number can classify each entity as an odd number or not an odd number.

**Notice that in this process the initial categorizing is accomplished by noticing something which is not really a property of the things being categorized but is instead a sound or mark artificially associated with the things. To what extent (if any) does this tend to dull one's sensitivity to the real basis of the categorizing, i.e., to the property or properties which would enable one to identify an instance of the concept when no linguistic clue is present?

***The vocabulary which he is capable of using in precise verbalizations of his own insights.

b) Why does attempting to formulate a statement of a universal generalization immediately after becoming aware of the generalization frequently

(1) result in the learner becoming confused about the generalization itself, and

(2) diminish his power to use the generalization in problem solving?

c) Would immediate linguistic formulation of one's discoveries have the same detrimental effects described in (a) and (b) above if the individual had sufficient linguistic power to avoid making incorrect trial sentences?

5. Can the dawning of awareness of existential generalizations and universal generalizations be identified in the learning of higher mammals other than man?

6. Can some kind of awareness in animal learning be identified with each of the logical connectives in language, thus establishing a non-linguistic basis for the penetrating flashes of insight in mathematical and scientific thinking--"flashes" for which a careful, logical formulation in language may require weeks or months of additional work?

7. At what age may a bright child comprehend that there is not a natural if-and-only-if relation between words of his mother tongue and the referents of those words, but instead that names are assigned to entities, the relation of symbol to referent being maintained (or destroyed) by those who use the words? (How can the "stage be set" in home or nursery school for a child to make this discovery as early as possible?)

8. For a given individual, how can we tell when it is not too soon for one to verbalize a discovery?

a) Should a newly-discovered entity be named as soon as the learner gives evidence (by classifying instances) that the awareness is there?

b) How soon after naming the entity should one be urged to attempt verbalizing non-redundant necessary and sufficient characteristics of the entity?

c) How soon after discovering a universal generalization of a conditional or a bi-conditional can one safely attempt linguistic formulation of it?

[Questions 7 and 8 are pertinent to identifying how soon deductive organization of knowledge is appropriate.]

9. If increasing one's certainty cannot be a motive for deductive proof until after one has acquired considerable skill in the process, then what is the primitive motive for this activity? Can it be the discovery that careful use of precise linguistic formulations yields prophetic power, a power of prediction less fallible than the unverbalized deductions usually classified as "hunches"? If so, how can more students be helped to acquire and appreciate this power of language without a taste for it being destroyed in the process?

10. Would it shed any light on the questions above to find out what kind of setting would impel an organism to use as a symbol an expressive sound or gesture previously accompanying the thing (or situation) to which the organism suddenly wanted to refer?

11. Can the different human motives for doing work be identified in animal learning? For example, do animals other than man engage in purposeful activity motivated by curiosity, as well as in activity for which the motivating drive seems to be

    a) material gain (food, shelter, provision for young, etc.)? or

    b) avoidance of unpleasantness, as in the case for many work horses? or

    c) the kind of "loyalty" or "faithfulness" with which a milkman's horse takes responsibility on his route, or with which a dog often assumes unassigned responsibility for guarding its owner's property? or

    d) the desire for another's approval, which motivates much of the learning of dogs in training? or

    d) the desire for self-approval--that is, achievement in accord with one's own philosophical beliefs?

    Among the rare trainers of top show horses, a strong curiosity and a long attention span in a very young foal are considered important prerequisites for the capacity to learn later to perform with the combination of brilliance, pride, courage, dependability, and natural animation which distinguish the occasional great show horse from other "good" performers. Can this combination of dependability, initiative, self-respect, and self-sustaining pride be associated not only with breeding, but also with

    a) certain features in the early handling of the individual animals (or the absence of certain features in early experience)?

    b) certain procedures in the basic training and finishing of great show horse individuals--procedures perhaps so subtle that even the trainer himself was unaware of them?

[Answers to questions 1, 2, and 4 in the preceding list are necessary before teachers (other than the few who discover those things for themselves, or who learn them by directed observation and in-service help as in the UICSM pilot schools) will believe it worthwhile to learn to do the kind of teaching through which students experience discoveries often enough to acquire a taste for research. Until these questions are answered, colleges will continue to train teachers who conduct schools in which children learn mathematics and science as interpretations of sentences written or spoken by others, a process that is drudgery in comparison to the exhilaration of re-discovery. Writers of textbooks and laboratory manuals will continue to present point-first all the "good stories" that might have emerged from directed investigation. And parents who were so taught naturally will start their own children in the same way—teaching them to chant, "one, two, three, four, five," etc., before matching sets of objects or events has ever made them aware of the entities for which the number words are names.]

Let us turn our attention now to questions 5, 6, and 10 on the list. Can it be that the whole deadly procedure prevalent in most classrooms may be based on mistaken notions of what language is, a mistake which not only leads to gross misuse of language in the promotion of learning, but which also detracts from man's power to use well his greatest-of-all-tools

a) to detach himself from unverbalized false beliefs,

b) to test his own beliefs (either by deductive organization or by search for counter-examples or verifications),

c) to communicate with precision, and

d) to record with correctness?

Is it possible that answers to questions 5, 6, and 10 in the list above might yield a break-through on this problem of what language is?

Finally, could comparison of training methods associated with the various motives under which animals can be taught to work cast any light on how education promotes or thwarts development of talent for, and devotion to, research? Furthermore, do the characteristics by which promising young individuals are selected by the most successful trainers of field trial dogs, seeing-eye dogs, and outstanding show horses cast any light on early identification of gifted human beings? Is capacity for successful development of such animals related to kinds of early experience with human beings, or to variety of experience during pre-training age? (For example, trifling with a foal, teasing it, playing with it without establishing both trust and obedience, can make the animal almost impossible to train later. On the other hand, the remarkable versatility and intelligence of the Arabian horse is often attributed to the constant and careful early handling of foals by the Arabs.)

TEACHING THE STRATEGY OF PROBLEM SOLVING

Paul C. Rosenbloom
University of Minnesota
and
Minnesota State Department of Education

This paper is based largely on my experiences during the past four years in teaching mathematics to students from the fifth grade up. This experience has occurred in various frameworks. I have held lectures for superior high school and junior high school students on Saturday mornings at the University of Minnesota. I have taught gifted fifth and sixth graders (IQ from 120 to 180) one day a week after school in the Minneapolis public schools. I have also taught demonstration classes from fifth grade up with unselected students on one short visit in more than 30 school systems in Minnesota and in several schools in other parts of the country. I shall also make a few remarks on teaching at the college level. It will be obvious to any mathematician that I owe a great debt to my own teachers, Professors Pólya and Szegö at Stanford University and Professor Rademacher at the University of Pennsylvania.

Before I go into details, let me briefly summarize what kind of behavior I wish to produce in the student. In the earliest stage the student should be able to solve problems which are presented to him. The difficulty will be graded, starting from direct one-step applications of previous knowledge and gradually proceeding to problems requiring analysis and restructuring. During this stage I perform the analysis of the problems and elicit the solutions by Socratic questioning. In an intermediate stage I wish the student to discover methods of solving problems. By a planned sequence of experiences I wish him to rise above the problem immediately before him and acquire a perspective toward families of problems. In the final stage I wish him to recognize problems himself, and to know what questions to ask himself. By this time he should have a sense of wonder and a taste for elegance.

In my classes of fifth and sixth graders I begin with the game of Nim. This is played by two people with toothpicks. The toothpicks are laid out in any number of piles and any number in each pile. Each player in his turn takes as many toothpicks as he wishes from any one pile at a time. The one who takes the last one wins.

After explaining the rules I have the children play against each other until they are familiar with the game. Then I play against the class, beating them every time, until they become aware that I have a strategy. I then present them with the research problem of finding the way to win at Nim. I tell the children that I won't tell them the solution, but I will show them how a research scientist goes about investigating such problems. I give them only the hint that they will need to know a new way of writing numbers.

During the next few (about ten) lessons I spend about two-thirds of each hour on binary arithmetic and the rest on their research on Nim.

Invariably someone notices during the first lesson that whenever I leave two equal piles I am sure to win, for I take from one pile whatever my opponent takes from the other. This leads to the concept of a winning position, one such that you can force a win no matter what your opponent does. We then begin recording the winning positions we have found, thus:

Winning Positions

$n \quad n$. (n any whole number)

Their first assignment is to play Nim at every opportunity and to record any winning positions they discover.

Not having had any experience in systematic investigation yet, the children may not make any progress at all at this time. Sometimes someone discovers that 1, 2, and 3 constitutes a winning position.

I now suggest that a scientist usually breaks up a big problem into smaller parts. He studies a simple case intensively at first. He varies the conditions systematically. I propose that they concentrate on the situation where there are only three piles and there is one in the first pile. I tell them to try small numbers in the second pile and see how many they must put in the third pile in order to have a winning position. For example, if there are 1 in the first pile and 3 in the second, then they must have 2 in the third pile, this being only a permutation of the known winning position 1, 2, and 3. This illustrates the common strategy of reducing to a previous case. The children now discover quickly the winning positions

1 4 5

1 6 7

1 8 9

They begin to discover a pattern. I encourage them to make conjectures by telling them Pólya's maxim. "It is better to have a wrong idea than no idea at all." A conjecture must be stated precisely enough that one can make a prediction which can be tested. The pupils soon arrive at the empirical generalization that

$$1 \qquad 2n \qquad 2n + 1$$

is a winning position. We then find an informal proof by mathematical induction.

The children then proceed to investigate the case where there are 2 in the first pile, and find the winning positions

2 4 6

2 5 7

2 8 10

2 9 11

and the generalizations

$$2 \quad 4n \quad 4n + 2$$

and

$$2 \quad 4n + 1 \quad 4n + 3$$

By the time they have found the position

$$3 \quad 4 \quad 7$$

someone usually makes the conjecture that the sum of the two smaller numbers is equal to the third. This is exploded by the counterexample

$$3 \quad 5 \quad 6$$

At this time we discuss the difference between empirical evidence and logical proof. I give the children further examples such as the proposition that

$$n^2 - n + 41 \text{ is a prime,}$$

which is true for $0 \leq n \leq 50$ and false for $n = 41$.

I now suggest to the children another way to arrange their results. We construct a table like a multiplication table:

	0	1	2	3	4	5	6	7	8	9	10
0	0	1	2	3	4	5	6	7	8	9	10
1	1	0	3	2	5	4	7	6	9	8	11
2	2	3	0	1	6	7	4	5	10	11	8
3	3	2	1	0	7	6	5	4			
4	4	5	6	7	0	1	2	3			
5	5	4	7	6	1	0	3	2			
6	6	7	4	5	2	3	0	1			
7	7	6	5	4	3	2	1	0			
8	8	9	10						0	1	2
9	9	8	11						1	0	
10	10	11	8						2		0

where in the m-th row and n-th column we place the number k which must be in the third pile in order that m, n, and k be a winning position. We observe that the winning position 1, 2, and 3 is recorded in six places in this table, corresponding to the permutations of the three piles.

The table makes evident other symmetries and patterns. It also makes the gaps in our knowledge quite obvious, and suggests a program for further research. I find it effective to record conjectures with a different colored chalk so that it is clear which are the known facts and which are guesses. By the next lesson the children have completed the above table.

By this time the children are familiar with binary notation. I tell them to record their results in this notation, thus:

$$3_{\text{ten}} = 11_{\text{two}}$$

$$5_{\text{ten}} = 101_{\text{two}}$$

$$6_{\text{ten}} = 110_{\text{two}}$$

The children are quite imitative and follow, without explicit instruction, this arrangement in which the units, twos, fours, etc. digits are written in columns. After making conjectures and some experimenting with four piles, the children now discover empirically the rule that a winning position is one in which there are, in the above arrangement, an even number of ones in each column.

In my experience, these children are not mature enough to appreciate a formal general proof. They do appreciate a proof presented in terms of a concrete case with reasoning which they recognize as perfectly general. Thus we may consider such a position as

$$7_{\text{ten}} = 111_{\text{two}}$$

$$12_{\text{ten}} = 1100_{\text{two}}$$

$$22_{\text{ten}} = 10110_{\text{two}}$$

$$29_{\text{ten}} = 11101_{\text{two}}$$

We ask what kind of change must your opponent make if you leave the toothpicks in this way? Look at the first digit, counting from the left, which is changed when he takes from one pile. In order for the number to decrease what change must he make? How will this affect the evenness or oddness of the number of 1's in that column? If you leave an even number of 1's in each column can your opponent leave such a position?

We also consider such a position as

$$7_{\text{ten}} = 111_{\text{two}}$$

$$18_{\text{ten}} = 10010_{\text{two}}$$

$$28_{\text{ten}} = 11100_{\text{two}}$$

$$29_{\text{ten}} = 11101_{\text{two}}$$

and question them until they discover the principle for reducing this position to a winning position (which can be done by taking 12 from the 18 pile, or 20 from the 28 pile or the 29 pile).

Details of the above procedure as applied to a number of topics can be found in my book, Adventures with Numbers, together with Teacher's Manual, published at present in a limited preliminary edition by Addison-Wesley Co., Reading, Mass.

A lesson which I give to average 9th graders on an isolated one-hour visit begins with having the students raise 1 + x to several higher powers. They compute $(1 + x)^n$ for n = 2, 3, and 4. I ask them how they can use their knowledge of $(1 + x)^n$ to compute $(1 + x)^{n+1}$. We discuss where the coefficient of any power x^k comes from, and from this the students discover the relation between successive rows in the table of coefficients:

n	1	x	x^2	x^3	x^4	x^5	x^6	x^7
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		
6	1	6	15	20	15	6	1	
7	1	7	21	35	35	21	7	1

I then ask them to predict the next three rows without actually multiplying out $(1 + x)^n$ and then to verify their prediction.

After some preliminary exploration of the table, I ask the class whether they notice anything about the 7th row. They reply that all but the two 1's are divisible by 7. I then ask whether a similar phenomenon occurs in any other rows. They say that it occurs in the 2nd, 3rd, and 5th rows. I then ask for a general rule. Some pupils think that 2 is an exception and that this phenomenon occurs in the odd numbered rows. Others conjecture that it occurs in the prime numbered rows—this often happens even when the students do not know the term "prime". The students are required to state their conjectures in a precise enough form that they can be tested.

I now write on the blackboard

$$(1 + x)^7 = 1 + 7S + {}^7,$$

where S stands for "something". We substitute x = 0 and x = 1 into the equation and check the relations

$$1^7 - 1 = 7S \qquad (S = 0)$$

$$2^7 - 2 = 7S \qquad (S = 18)$$

We substitute x = 2 and apply the last result (using subscripts to distinguish the various "somethings"), and obtain

$$\begin{aligned} 3^7 &= 1 + 7S_1 + 2^7 \\ &= 1 + 7S_1 + 2 + 7S_2 \\ &= 3 + 7(S_1 + S_2). \end{aligned}$$

$$3^7 - 3 = 7S_3 ,$$

which we also check.

I then write, "Suppose $x^7 - x = 7S_1$, i.e. $x^7 - x$ is divisible by 7. Then

$$(1 + x)^7 = 1 + 7S + x^7$$

$$= 1 + 7S + x + 7S_1 ,$$

and

$$(1 + x)^7 - (1 + x) = 7(S + S_1) = 7S_2 ,$$

so that

$$(1 + x)^7 - (1 + x) \text{ is divisible by 7."}$$

I then rewrite just the hypothesis and the conclusion thus:

"If $x^7 - x$ is divisible by 7 then $(1 + x)^7 - (1 + x)$ is divisible by 7."

Now I ask someone in the class to read this statement with "x" replaced by 3. I ask the class to tell me what conclusion they draw from this "if-then" statement combined with the fact that $3^7 - 3$ is divisible by 7. We repeat the process with $x = 4$.

Then I tell the class to imagine a row of tin soldiers placed so close to each other that if anyone is tipped over, then the next one falls down. What happens if we tip over the first tin soldier? What conclusion do they draw from the "if-then" statement, together with the fact that $1^7 - 1$ is divisible by 7?

We then discuss the fact that the only relation used in the above reasoning is the phenomenon observed in the 7th row of Pascal's triangle. What conclusion would they obtain by applying the same argument to the 2nd, 3rd, and 5th rows? If their conjecture about the general law is true, then what is the general consequence?

In less than an hour the class itself gives me the statement that

If p is a prime and x is any whole number, then
$x^p - x$ is divisible by p.

In teaching the concept of integration I find it effective to start with Newton's laws of motion and lead to the problem of solving differential equations. I then suggest the technique of the scientist of beginning with the simple class of equations of the form

$$y' = f(x).$$

Before we attack this problem in detail, I propose to the students the questions which a research mathematician would ask himself when confronted with a new problem of this kind:

(a) Does a solution exist?

(b) Is the solution unique? If not, what data are needed in order to determine the solution uniquely?

(c) How does the solution depend on the data? If the data are known to within a small error, is the solution determined to within a small error?

(d) Are there any simple cases which we can solve explicitly?

(e) How can we compute the solution practically?

These questions suggest a research program. The questions are not, of course, answered in this order. If the further questioning is properly planned, the students discover for themselves, by quite natural steps, the Perron integral. The Riemann integral arises as an answer to the last question.

The books of Pólya give a further detailed exposition of heuristics and strategy, with many concrete examples. The collection of problems by Pólya and Szegö, Aufgabenund Lehrsätze aus der Analysis (Springer, 1925), has a wonderful preface which gives many valuable precepts for approaching problems. The problems in this book are so arranged as to lead the student to discover several branches of mathematics for himself.

If we measure a teacher's ability by the number of outstanding students he has produced, then the outstanding mathematics teacher in this country is R. L. Moore, of the University of Texas. His technique is to present the students with the postulates, definitions, and theorems in a well-planned order, and to have them develop the subject themselves. He even goes to the extreme of removing all books on the subject from the library. It should be noted that, although the subject matter covered in his courses is quite narrow, his students frequently branch out into quite different fields after they leave him and apply their creativity to domains which he has not taught them. Wilder and Moise are conspicuous examples of this. We should also emphasize that Moore does not try to cover any particular amount of material in a given time, but puts the main effort on developing the students' initiative and independence.

At present we now have scientific evaluation of the program which I started in the elementary school. This year Mr. Marshall Kamer, of the Calhoun School in Minneapolis, has taken over the experimental class and is conducting a more careful experiment. In this effort he has the assistance and cooperation of Dr. Arthur J. Lewis, Assistant Superintendent in charge of elementary education, and his staff. We hope, in the near future to carry out an experiment meeting the most rigorous scientific standards for determining the effect of such instruction on problem solving ability.

THE MINNESOTA NATIONAL LABORATORY
FOR THE IMPROVEMENT OF SECONDARY SCHOOL MATHEMATICS

by

Paul C. Rosenbloom
University of Minnesota
&
Minnesota State Department of Education

This is a brief description of an agency for education experimentation which is being established within the Minnesota State Department of Education.

The Minnesota National Laboratory for the Improvement of Secondary School Mathematics is established as part of the division of instruction in the State Department of Education, and is under the jurisdiction of Farley D. Bright, Assistant Commissioner of Education in Charge of Instruction. The Laboratory has a small central staff, to be supported by the State of Minnesota, consisting of myself as Director, an assistant to be appointed, and a secretary. I have an advisory committee consisting of a statistician, a psychologist, mathematicians from the University, and state and liberal arts colleges, a professor of education, and representatives of school administrations, teacher organizations, business, and industry.

At the beginning of September application forms for participation by teacher and by school were sent to every superintendent in the state. A crude measure of teacher qualification will be set up, the population of qualified applicants will be stratified according to the population of the community, and teachers will be selected by random sampling from each stratum.

Participating teachers will have the responsibility of studying each unit in preparation for teaching it, administering tests to be sent out, writing reports, attending meetings and seminars, and keeping certain records.

Coordinated with the Laboratory, but not officially connected with it, is the Minnesota School Mathematics Center, also directed by me, and affiliated with the Minnesota Academy of Science. We have a 7th and 8th grade writing team which continues the work done at Yale last summer, and are now organizing a 9th grade writing team. An experimental class has been placed at each school where there is a member of the writing team. By this close association of the writing team with experimental teachers we obtain a quick feedback of information.

We expect that both the Laboratory and the Center will be permanent institutions. If the Laboratory should prove successful, the State Board of Education contemplates extending it to all other fields as the corresponding national projects for curricular revision are organized. We hope in this way to establish research in curriculum as a normal and essential activity of the State Department of Education. We hope that our experimentation program in mathematics will serve as a pattern for other states and for other subjects. In this way we may contribute to a general improvement in the quality of education in the United States.

A RATIONALE FOR DIRECT OBSERVATION OF VERBAL BEHAVIORS IN THE MATHEMATICS CLASSROOM

E. Muriel J. Wright
Washington University

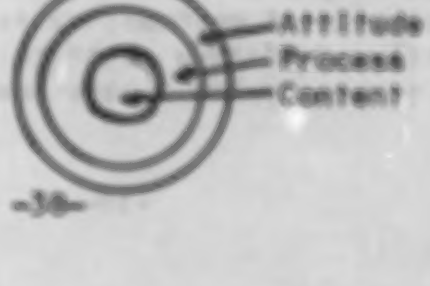
As a primary arena for learning, the classroom its participants and activities, and the results of specific learnings, have long been an object of study. On the whole, systematic study has emphasized indirect assessment of the classroom variables, for instance, study of teacher qualification—academic, professional attitude, etc., and of pupil preparation—achievement and change in achievement, attitude, etc. Such indirect assessment—though permitting refined measurement technique—by-passes the arena itself and requires itself such assumptions as: stated teacher attitudes are actually those demonstrated in the classroom; or pupil change in achievement is the direct result of particular experiences in a specific classroom. With the gradually improving techniques of systematic observation of behavior, many recent studies of classroom learning have sought the advantages of direct assessment of the behaviors of teacher and pupils....the most fruitful being concerned with consideration of verbal responses.

Yet almost without exception these studies have considered behaviors of teachers and pupils in important yet peripheral aspects of the classroom: those which set the stage, so to speak, for successful mastery of subject matter. Believing that such peripheral aspects as social climate, etc., generally may be viewed as a function of the presentation of subject matter, I have directed my attention to close study simultaneously of the content, the process and the attitude contained and inferred in the thinking verbally expressed. Specific definitions of the categories defining the areas of content, process and attitude proved at the same time to be characteristics of discursive thinking, and to be generally accepted aims of mathematics teaching.

In this paper I wish to support the study of classroom verbal behaviors by these broad areas, and to outline a method for employment of their specific definition as a system of categories for direct observation of the mathematics classroom.

I. STUDY OF CLASSROOM VERBAL BEHAVIOR BY AREAS OF CONTENT, PROCESS, ATTITUDE

Understanding of the meaning of the language used in mathematical argument necessarily involves two aspects: mathematical content and psychological process. An additional factor must be considered in a learning situation—that of, if you will, sociological attitude: for the cognitive aspects have resulted from or have been effected by impetus from the general or particular environment of the speaker. I have found it convenient to present these three aspects in the study of language schematically by three concentric circles: the innermost area, mathematical content; the middle area, psychological process; and the outer area, sociological attitude.



This schema encompasses the inference that the environment has induced certain processes in the consideration of particular subject matter. This schema also is in marked agreement with Piaget's (1) analysis of logical thinking where he points out the affective and cognitive aspects of thinking—his cognitive aspect here divided into process and content. This division of the cognitive aspect is in keeping with Cassirer's (2) consideration of thought as "thing-concepts" and "relation-concepts" with the science presenting the things as distinct and certain, and the psychology as presenting the relations as a temporal process. "Psychology in this sense gives the approach to problems, which must seek their (complete) solution in logic and in their application to science".

The ordering of these concentric circles also may be commented on: the outer area, attitude, being larger than the process permits inclusion of discursive thinking and other types of thought and considers the many specific attitudes that may be affective in different types of thinking; the center area, process, being larger than the content implies application of the process in many subject areas. Viewing the circles from above brings out the importance of simultaneous consideration of the three areas for thorough assessment of the discussion of particular subject matter, yet implies that certain attitudes and certain processes may be pertinent only to a particular subject field.

While the language of the classroom—including its mathematical refinements and its approximations by gesture, hierogliphics or drawings—may be studied from these three reference areas, the problem should be raised of the place of this verbal behavior as a sampling of the thinking actually occurring in the speaker. Langer(3) points out that "language has a form which requires us to string out our ideas even though their objects rest one within the other" and that this implies the limitation on language that only thoughts that can be arranged in such an order can be spoken at all. It seems important at this stage to avoid such assumptions that language is the only means of articulating thought and that everything which is not speakable thought is feeling. Such assumptions are surely inadmissible in the light of the mind's organized storing of the meaningless impressions of relatively simple sense experience or in the light of the complex symbolism of the musician and the painter. Langer suggests the term presentational thinking as complementary to discursive thinking and as descriptive of thinking which cannot readily be ordered temporally and therefore cannot be expressed in language. However, while thought itself must include more than discursive thinking, mastery of mathematical subject matter itself implies an understanding of specific mathematical ideas which finally may be demonstrable by the discursive symbolism of language. It is concluded that the verbal behaviors of the classroom provide sound materials for an approach to the study of thinking about or learning of mathematical problems.

II. EMPLOYMENT OF THESE AREAS FOR DIRECT OBSERVATION OF THE MATHEMATICS CLASSROOM

Turning now to consideration of the study of content, process and attitude in the verbal behaviors of teacher and pupils of a mathematics classroom. Expressed in terms of aims of mathematics teaching, we may consider appreciation of content, skill in strategy of thinking, and certain desirable attitudes. For a period of about three years I have been defining and refining systems of categories to express these three aims specifically, and to fulfill the observational requirements of exclusiveness, of exhaustiveness and of unity by repeated use in systematic observation of mathematics classrooms. The aims and the third experimental form of the systems of categories are outlined in Tables 1, 2 and 3. With some revisions necessitated by current observations the categories are:

Appreciation of mathematics: fundamentals (math. elements and logic)
relationships
applications to mathematics
applications to other areas
techniques (addition)

Ability to think: analyzing
synthesizing
specializing
generalizing
Relevant information

Attitude of curiosity and initiative:
curiosity
independence
receptivity

Neutral

The categories of appreciation and attitude are, on the whole, simply related to their names; however, a brief statement of the definition used for the categories of the frame ability to think may be useful. Analyzing occurs in working back from an assumption of a desired conclusion to accepted principles; synthesizing - its opposite—occurs in working from accepted principles towards a desired conclusion. These two approaches provide the successful method of attack on most mathematical problems, and together might be described loosely as the deductive method. An examination of the method of proof of a particular problem will show that a combination and ordering of the steps used in analyzing and synthesizing have actually provided the demonstration of the solution of the problem. Specializing contains two important aspects—assessment of analogous problems, and application of a generalization in a particular problem. Opposing this is generalization: the recognition of significant attributes of a number of specific instances, and also the passing from these instances to understanding of a common characteristic. Together these categories correspond loosely to the inductive method. A fifth category was found necessary to include the behaviors that did not contribute definite direction to the problem at hand: analyzing and synthesizing contributed motion to or from the desired conclusion; specializing and generalizing contributed motion to a wider or a narrower problem. The fifth, "relevant information", encompassed mathematical ideas potentially valuable for accommodation in the other more active categories. The present understanding of this category as that of psychological study of process in thinking has suggested some value in renaming of the categories perhaps in correspondence to certain of Duncker's (4) division:

e.g. analyzing — organic
synthesizing — mechanical
specializing and generalizing — restructuring

During the periods of observation in the classroom, behaviors are viewed in each of the three frames and are classified once under each. The teacher might suggest: "Let us list the steps used in attaching these problems": illustrating (content) mathematical application, (process) generalizing, and (attitude) encouragement of independence. The student might estimate an answer: "Bill's answer seems way out—the angle is just less than 45 degrees so the tangent function would be about point nine (.9)": illustrating (content) fundamentals—estimating and knowledge of tan 45—, (process) synthesizing, (attitude) independence. The present classification is quite pleasing in permitting classification of such a response—that of estimating an answer or making a plausible guess—which lack had indicated fault in an earlier set of categories. It is found that all behaviors relating to the lesson can be classified in each of the frames. Others not concerned with the subject at hand, are classified as Neutral. Examples might be: "Please shut the door", "Hand in your homework". Other specific examples appear in tables 1, 2 and 3.

Single behavior

The single behavior is obtained by time sampling: the behavior of the first speaker in a fifteen second interval. During each minute of observation, two behaviors are classified. Of the four fifteen-second intervals occurring, the first and third are used for observation, and the second and fourth for classification.

Unit of Behavior

The single lesson provides a natural unit for observation of the classroom. Despite the variation in the length of periods from one lesson to another, no fixed length of time for observation within the lesson has proved feasible. Therefore all verbal interaction occurring while the teacher and pupils are working as a group is classified. The single lesson period is usually about fifty minutes in length resulting in classification of about one hundred behaviors. If the purpose of the observation is that of obtaining a representative sample of behaviors of a classroom for comparison with others, behaviors from ten class periods selected either in order or at random have been found to provide sound samples of total classroom discussion.

The observer

The observer views and classifies behavior from a desk at the back of the classroom using duplicated recording sheets and a stop-watch. The classification or behaviors under the defined categories is determined by the understanding, interpretation and inference of the observers. Care, therefore, must be directed to adequate preparation in terms of mathematical understanding generally, and in terms of the particular topics observed as well as in terms of skill in the practical procedures.

III. STUDIES TO DATE AND LIMITATIONS OF INSTRUMENT

Most investigations carried out to date have been with a view to establishing the validity of the categories or the observational procedures. Several revisions of the categories have occurred permitting sharper classification and finer agreement between observers. Various approaches have supported the choice of the fifteen-second division in observations, and of the ten-period visit to provide representative results of a classroom as sampling procedures. Finally means of representing results for clear description and comparison of classrooms are developing and undergoing refinement. So far attention has been directed to frequencies or proportions of behaviors in individual categories, rather than seeking over-all indices of Content, Process and Attitude. This is partly because of limitations imposed by the nature of the data and the small frequencies in many single categories, and partly because of the clearer meaning when the categories are viewed singly or in combination yet not absorbed into an over-all index. (X^2 is suitable in terms of characteristics of variables but unsuitable in light of characteristically small frequencies in certain categories.

Two areas require further assessment in the use of the instrument: the treatment of the data mentioned above; and the preparation of observers. The latter is now limited by the need of training by myself or by observers trained by me. Studies currently underway and others planned should contribute to removal of these limitations.

IV. AREAS FOR INVESTIGATION WITH SUCH AN INSTRUMENT

There are several ways in which this instrument might be profitably used. One of its greatest interests to me personally has been the marked sharpening of my attention to details in the content and process area discussions in mathematics lessons. Already a similar heightening of attention to the several aspects of the three frames, and essential considerations of the lesson has been apparent in student teachers making observations with it.

It has become apparent also that by redefinition of exact categories in the three frames to meet the peculiar needs of another subject area, comparison of categories common to two or more subject areas would be possible. It is likely that in such a redefinition the area of process would prove the most stable.

Now within the mathematics classroom many problems for fruitful study may be suggested: generally the results would be expressed in terms of patterns of categories within the three frames content, process, attitude. Among these problems are:

i. The relationship between classroom behaviors and individual pupil behaviors in problem solving.
ii. The relationship between anticipated and observed behaviors in the mathematics classroom
iii. The variation of teacher behaviors while teaching the same topic to several groups of pupils.
iv. The effect of various broad areas of subject matter and method on patterns of behavior in the mathematics classroom.
v. The stability of patterns of behaviors in the mathematics classroom throughout the solution of a closed problem, an open problem, a topic, a semester.........

Finally, in the light of my initial concern with limitations of the indirect assessment of the classroom variables, comparison should be carried out of results of this systematic observation as direct study and methods of indirect consideration of the mathematics classroom.

REFERENCES

1. Piaget, Jean, *The psychology of intelligence*, Routledge, London: 1950, p. 4.

2. Cassirer, Ernst, *Substance and function*, Dover Publications, New York: 1953, pp. 9, 343.

3. Langer, Susanne K., *Philosophy in a new key*, New American Library, New York: 1948, pp. 63, 70, 73, 79.

4. Duncker, Karl, On problem-solving, *Psychological Monographs*, *58*, 5, whole no. 270, 1945, p. 45.

TABLE I: Definition of the several categories of frame of reference, A, Process: ability to think, and their positive and negative qualities. Particular example and symbolic classification of a behavior in each category.

Frame	Category and Definition	Example	Symbol
Process: ability to think, A.	1. analyzing Working from an assumption of a desired conclusion towards accepted principles.	*(The problem--to show the roots of a particular equation are real and unequal) Pupil: If the roots of $16x^2-16x-21 = 0$ real and unequal, then we should find the discriminant greater than zero	$A_{1+}(B_{3+},C_{2+})$
	2. synthesizing Working from accepted principles towards a desired conclusion.	Teacher: We have decided that the three parts of the whole dividend may be expressed as \$2x, \$3x, \$1x. We also know that this whole dividend amounted to \$66. What may we conclude from these two statements?	$A_{2+}(B_{4+},C_{2+})$
	3. specializing The use of significant attributes of a given set in an analogous set or the applying of a given set to a smaller included set.	Pupil: This is a geometric series, so to sum "n" terms I'd use the formula $S = \frac{a(1-r^n)}{1-r}$	$A_{3+}(B_{3+},C_{2+})$
	4. generalizing The recognition of significant attributes of a given set and the passing from the consideration of the given set to that of a larger inclusive set.	Pupil: That answer can't be right, it doesn't agree with the answer in the back of the book.	$A_{4-}(B_{\overset{3-}{4-}}, C_{2-})$
	5. relevant information Statement of information relevent to the mathematics lesson but not contributing to forward movement of the lesson.	Teacher: Bill, will you read the first problem aloud. Pupil: My answer to question is $x^2 - y^2$.	$A_{5+}(B_{3+},C_{2+})$ $A_{5+}(B_{3+},C_{3+})$

TABLE 2: Definition of the several categories of frame of reference, B, Content: appreciation of mathematics, and their positive and negative qualities. Particular example and symbolic classification of a behavior in each category.

Frame	Category and Definition	Example	Symbol
Content: appreciation of mathematical structure, B	1. fundamentals Fundamental elements, operations, postulates, method of proof.	Pupil: Of the three alternatives: y 1 y = 1 y 1 we must show the impossibility of the first two in order to prove that y must be less than 1.	$B_1.(A_1.C_2.)$
	2. relations Theory derived from fundamentals.	Teacher: Now we have been able to show that 2 = a/b where a and b are different integers prime to each other.	$B_2.(A_3.C_3.)$
	3. mathematical application Theory applied to mathematical examples.	*(The problem--to show that the roots of a particular equation are real and unequal) Pupil: If the roots of $16x^2-16x-31 = 0$ are real and unequal, then we should find the discriminant greater than zero.	$B_3.(A_1.C_2.)$
	4. other application Application and importance of mathematics to problems from other fields, for example, physics, chemistry, economics, sociology, history.	Teacher: Now $s = 1/2gt^2$ is a particular formula used in physics. The t represents time, s, distance and g the effect of gravity. This function of t describes the distance a body falling freely under the force of gravity will move in relationship to the time passed.	$B_4.(A_3.C_1.)$
	5. technique Use of materials of tools without attention to mathematical significance.	Pupil: (Using slide rule with ease) The answer is 3.82.	$A_3(B_5C_2.$ or $C_3.)$

TABLE 3: Definition of the several categories of frame of reference, C, Attitudes: curiosity and initiative, and their positive and negative qualities. Particular example and symbolic classification of a behavior in each category.

Frame	Category and Definition	Example	Symbol
Attitudes: curiosity and initiative	1. Enthusiasm Enthusiasm for fresh knowledge. Arousing of active interest.	Pupil: But that is a chemistry problem!	$C_{1}.(A_{3}.P_{6}.)$
	2. Independence Independence of thought and action including perseverance and self-directed work towards the desired outcome.	*(The problem—to show that the roots of a particular equation are real and unequal) Pupil: If the roots of $16x^2-16x-31 = 0$ are real and unequal, then we should find the discriminant greater than zero.	$C_{2}.(A_{1}.P_{3}.)$
	3. Receptivity Teacher: Sustaining of passive interest e.g. during teacher demonstration.	Teacher: Now the next step should be clear: if $24\ p^2 = 8q$ then $3p^2 = q$	$C_{3}.(A_{2}.P_{3}.)$
	Pupil: Directed work towards desired outcome, e.g. direct response to question.	Pupil: My answer to that question is 3x.	$C_{3}.(A_{3}.P_{3}.)$
Neutral	No apparent effect on achievement of the defined aims. Behaviors that cannot be readily classified under at least two of the above frames. These are usually related to matters of class organization or discipline.	Teacher: Pass in your homework questions. Pupil: May I sharpen my pencil?	N

Omitted behaviors are noted when, during a fifteen second interval, no verbal behavior or blackboard demonstration subsequent to an earlier behavior is in progress. Supervised study is also noted to ensure exhaustive classification of the time period. Attention is drawn to the starred example which occurs under each of the three frames to emphasize that each behavior is necessarily viewed in each of the frames.

A METHOD OF DETERMINING THE CORRELATES OF CHANGE IN SKILL IN MATHEMATICS*

Philip H. DuBois
Washington University

A basic problem in research on learning is how to measure change in human behavior. We need to measure the same function at two points in time and to derive a joint function of these two measurements that will represent the change that has taken place from the first point to the second. The actual execution of this manipulation is more difficult than it seems at first.

Obviously an adequate solution to the measurement of change depends upon an adequate solution to the measurement of functions at single points in time. However, it is quite possible that we might be able to obtain measurements that, something like a snapshot, show status at one particular time without our knowing how to combine these measurements with later measurements so as properly to reflect movement or growth or gain.

Let us consider some of the methods by which learning has been measured in psychological experiments during the past half century. One of the oldest methods has been to find tasks in which proficiency at the beginning of the investigation could be taken as zero. It then seems reasonable that any proficiency demonstrated later on can be taken as a measure of gain. This procedure is perhaps not unreasonable when we confront an untrained rat with a maze, or a man with a list of nonsense syllables. It is unreasonable in a situation in which beginning students have widely varying knowledge and differing degrees of skills in the area in which they are to be trained. There is probably no important learning situation which a human adult enters with zero previous attainment. Even in learning nonsense syllables, in which there may be no prior knowledge of the particular list, the person may have considerable knowledge of the best ways to handle the task. For any organism, the course of learning is conditioned not only by the nature of relatively permanent structures but also by changes in those structures resulting from earlier experiences in adjusting to the environment.

In learning situations in which it is obvious that all subjects do not start at the same point, use has been made of the simple arithmetic differences between the measurement representing final status and measurement on the same scale representing initial status. In some of the simpler learning situations physical ratio scales are used to dimensionalize behavior. We can call to mind scores in terms of number of errors, circular error, time on target, time to complete a standard task, distance covered, and the like. In a highly abstract sense, measures of errors, time, weight and distance certainly fall on scales with equal units and with true zero points. Nevertheless when we apply these measures in a specific situation their mathematical characteristics do not necessarily continue to represent the real

---

*Prepared under Contract ONR 816(02) between Washington University and the Office of Naval Research.

situation. The reason is perhaps that only in very limited or abstract situations are the units truly equal. As the number of varistes involved in a phenomenon increases, the measurements yielded by the scales change. For example, as long as I am dealing with wire of defined composition and diameter, lengths are additive and, of course, subtractive. When we organize the material more complexly the situation changes. The problem of building a tower 200 feet high is not twice the problem of building a tower 100 feet high. In materials used, in effort expended, and in cost, the problem of the 200 foot tower may be several times as great as the problem of the 100 foot tower. The only dimension in which one tower is exactly twice the other is a highly abstract dimension of incidental importance.

So it is, I think, when we measure learning with scales borrowed from the physical sciences. When a rat progresses from 24 errors to 12 errors in a maze, it is only in a highly abstract sense that it has gone half the distance from its starting point to complete mastery. In terms of metabolic activities required, or time needed, either of which might be taken as measures between the initial level and the final level, the point at which the errors were reduced 50% might not be the half-way point.

Estimates of "percent gain," which are sometimes used as a measure of learning, assume that the researcher is using a ratio scale. It may be pointed out, however, that the use of scales with additive units and known zero points, such as errors or seconds or units of work accomplished, do not necessarily guarantee that sets of differences between initial and final status constitute meaningful variables that can be used in investigations of conditions influencing learning. The trouble with a crude difference score is that it is a mere positive weighting of final score and negative weighting of initial score, with effective weights completely unknown. Its correlations with variables describing the learners and defining the external conditions of the study, are greatly influenced by the vagaries of its composition. That crude gain will correlate negatively with initial status is practically certain; the degree of this correlation is a function of the standard deviations of the measures reflecting initial and final status, as well as the correlation between them. Our first problem, then, is to develop an appropriate measure of gain or change in function which can be used in studies of the learning situation.

Let us consider two points in time at which we have adequate and reliable measures of some aspect of behavior. Let z_1 be our first measure in z-score form and z_2 our second measure. We can divide z_2 into two uncorrelated portions: z_2 which is perfectly correlated with z_1, and a residual score, $z_{2\cdot 1}$, which is uncorrelated with z_1.

It is proposed that in learning situations in which all subjects do not start at a common point that this residual, $z_{2\cdot 1}$, be used as the measure of change. The use of a residual score to define change has several advantages:

1. The residual can be defined precisely and meaningfully.

2. The residual in z-form is completely independent of units of measurement and hence is applicable even if initial score and final score are in different metrics.

3. The residual is completely independent of and completely uncorrelated with initial status. Hence when it is correlated with outside variables, the correlation is with gain defined as independent of initial status, rather than a mixed function, an unknown part of which is initial status.

4. The residual is a measure of learning that does not require the use of a ratio scale to measure initial and final status. By the use of a residual score the question as to whether those who know most at the start of the study learn more than the others cannot be answered, since the answer to such a question would depend on the use of the same ratio scale at the beginning and end of the study. However, there is no restriction on investigations of outside correlates of learning.

5. The residual score as a measure of learning fits perfectly into correlational statistics in terms of variances. The variance of the final score is divisible into two uncorrelated portions: a part associated with and a part independent of initial status. The independent part can in turn be analyzed to find proportions associated with outside predictors of learning. These analyses should help us understand what goes on during the learning process.

6. While only first order residuals are considered here, it is perfectly feasible statistically to use residuals of any order as measures of learning, thus dissociating gain from any number of variables known to be associated with final status.

In a quantitative investigation of learning, three categories of variables may be distinguished: 1) the dependent variable; 2) independent variables; and 3) control variables.

Within a defined learning situation, it is possible to identify four principal sources of variance of the proficiency of individuals at any stage in their training. These are:

Prior skills and knowledge.
Aptitudes and capacities of the learner.
Goals, motivation and readiness to learn.
Effect of activity, or reinforcement.

In addition, for purposes of analysis, situational factors can be considered as variables.

In investigating learning by means of residuals, we remove from the measure of proficiency at one point in time the variance predictable from proficiency at an earlier stage. The resultant measure is a meaningful dependent variable which can be related by multivariate correlational analysis to independent variates, singly or in combination.

Theoretically, the number of control variables is unlimited. The appropriate procedure is to remove from the independent and dependent variables, prior to determining the net relationship, variance predictable from the control variables.

Experimental controls and statistical controls are theoretical equivalents. Both have the same purpose, namely, the reduction of the correlations of certain extraneous variables to zero. We are using various types of correlation to unravel the complex of variables in on-going situations, and to see whether or not we can find essential relationships between the pairs of variates under scrutiny. When we are dealing with a large number of variables a rather arbitrary decision sometimes has to be made as to where in our setup we are going to put any given variable, that is, whether we will treat it as a predictor or whether we will use it as a variable to be controlled.

Several types of correlation of residuals, including partial correlation, part correlation, multiple partial correlation, and various varieties of multiple part correlation afford complete flexibility in modifying one or more of experimental variates so that the correlation with any of the outside variates becomes zero. Provided that the variables are measured in such a form that they are amenable to correlational analysis, the method is simply to subtract from a variable that portion of its variance which is associated with one or more outside variables. Thus we can estimate the correlation between any pair of the experimental variates with one or more of the concomitant variates controlled. In deciding on the correlational model to apply in a learning study, the important question is always which type of correlation will yield the most accurate picture of the relationship in which we are interested.

Closely allied to correlational analysis is the analysis of covariance. Both employ identical logic and identical basic formulas. Both reduce relationships to zero by subtracting variance associated with primary variates. However, analysis of covariance uses equating of groups (which is an experimental control) as an additional means of holding some of the correlations at zero. Analysis of covariance also includes methods of treating nominal or unscaled variables. When our variates are based on the scales that are appropriate for computing correlations, and all relationships are linear, there should be really no difference in conclusions whether correlational analysis or analysis of covariance is used.

The method I have been describing is, of course, perfectly general, rather than specific to research in mathematics training. It would appear, however, that effective research on mathematics training must, in good measure, be carried out in on-going situations where only a limited degree of experimental control is feasible. To achieve dependable knowledge in this area, it seems that it is necessary to perform studies in which chief reliance must be put on statistical rather than experimental controls. It is also likely that we must replicate these studies until we are reasonably sure that our results are dependable. Fortunately, the scaling of proficiency measures in mathematics is perhaps a less formidable task than in areas where proficiency cannot be gauged with the student seated at a desk with paper and pencil at hand.

SCALING METHODS APPROPRIATE FOR DEVELOPING MEASURES OF ACHIEVEMENT FOR RESEARCH ON MATHEMATICS TRAINING

Goldine C. Gleser
University of Cincinnati Medical School
and University of Illinois

In order to carry out a program of research aimed at improving the quality and scope of mathematical knowledge among the children of our nation, it is of prime importance that we have satisfactory methods of evaluating the results of our efforts. No matter how diligent these efforts, how ingenious our teaching methods, how sophisticated our research designs, we will be seriously hampered in obtaining optimum results unless we have developed measuring instruments which are adequate for evaluating the resulting amount and rate of growth in performance.

In the past forty years great strides have been made in the field of psychological and educational measurement and much of this knowledge has been applied to the construction of standardized achievement tests. Many of these tests can be used to measure achievement for various purposes. However, these tests have primarily been designed in accordance with the aims of the traditional mathematics curriculum with its emphasis on skills in mathematical operations, definition of terms, etc., rather than on an understanding of principles and relations and on the ability to use them inductively and deductively in problem solving situations. Thus, where new methods of teaching and motivation are used with the aim of achieving these latter goals, it is necessary to devise new instruments for measuring the outcome of such treatments. Some of the important requirements of these tests are: that they be reliable and valid samples of the desired outcomes and that they sample this domain of outcomes along relatively narrow, independent dimensions which can give meaningful measures of achievement in important aspects of the desired behavior.

The first major requirement for any measure is that it be valid. In the case of achievement tests validity can only be assessed in terms of the educational objectives in the field in question. An achievement test is valid if it is a representative sample of the skills, knowledge and understanding which are considered important by the educators and around which the curriculum or program of studies is developed. Thus, the first step in developing a measure of educational achievement is the careful definition of the outcomes which the educational program is intended to accomplish. This implies not only a statement of the broad general goals of a course of studies such as mathematics, but the delineation of specific identifiable behaviors which it is desired to inculcate. If teaching is aimed at enabling the student to tackle new problems or more advanced work with understanding, then this should be specified. If it is sought to develop skill in reasoning and critical thinking, consideration should be given to the types of situations where such skills can be expected to be demonstrated. Furthermore, since what is acquired depends to a considerable extent on the emphasis and methods of teaching employed, it is important that these goals be broad enough to include all the relevent desirable outcomes to be expected under the various methods of teaching being considered, so that comparisons among them are possible. For example, the "discovery" method of teaching may perhaps lead to greater acquisition of skills in inductive reasoning as compared to deductive reasoning, whereas, presenting generalizations and their application may do the reverse. However, items involving both deductive and inductive reasoning are needed in an achievement test if in the opinion of educators both types of skills are important goals of training in mathematics and it is desired to know which teaching method best achieves both results.

The importance of an adequate consideration of different possible outcomes in the construction of achievement tests for the evaluation of teaching methods is aptly illustrated by a study of Brownell and Moser (1) on third grade arithmetic. They utilized two methods of teaching subtraction: 1) "decomposition" or borrowing one from the tens position in the minuend; and 2) "equal additions" or adding one to the tens position in the subtrahend to compensate for the ten added to the unit position of the minuend. A further differentiation was made in that some classes using each method were taught with emphasis on meaning while others were taught the method mechanically. The methods gave almost equivalent results when accuracy at the end of training and retention six weeks later were used as criteria. Either method, taught meaningfully, had a slight advantage. However, when a measure of performance was used involving transfer of training from 2-digit numbers to 3-digit numbers, it was found that classes taught by the "borrowing" method with emphasis on meaning had a decided advantage over all the other classes.

The first essentials, then, for designing achievement tests are: thorough knowledge of all the important goals of the educational program and skill in translating these goals into representative samples of behavior which can be scored objectively. This involves not only adequate representation of the areas of knowledge but also representation of the ways in which that knowledge can be expected to be applied. For example, in the field of simultaneous equations, the ability to recognize that a pair of values satisfies two simultaneous conditions is not the same as determining what values satisfy the two conditions, nor can it be inferred that the student who can perform the one task can necessarily perform the other. The ability to multiply two binomials does not necessarily indicate that the principle of algebraic multiplication is understood. Skill in writing test items which adequately sample the desired behaviors is extremely important to test construction.

A consideration of the many goals for a curriculum in mathematics or any other educational field immediately indicates that the measurement of achievement in these fields is not a unidimensional concept. The individual is hopefully advancing along a broad front of knowledge, skills and understanding, involving many conceivable aspects. In order to measure growth or achievement, however, we must identify factors or dimensions within this domain which are relatively independent of each other. My point may become more clear if we consider the example of physical development. If we desired to know to what extent children had grown physically in a given period of time, one would immediately ask "grown in what respect?" Physical growth can be measured in many ways, some more or less independent. Thus, we might determine height, weight, chest girth, hip girth, lung capacity, weight lifting, endurance, etc., at the beginning and end of the period in question. These variables are not all independent dimensions of physical development, but they are indications of different aspects of that development. Several of them overlap each other to a considerable extent. Weight tends to overlap height and chest girth, although it also depends on the relative amount of fat, bone, and muscle tissue and therefore may be somewhat related to strength. The children who had increased most in stature might also show a substantial increase in weight but this would not be completely independent information. If we used all these measures separately or combined them in some fashion to obtain an over-all index of physical development, the index would be affected to an unknown extent by the stature factor. Furthermore, another set of physical measures with different overlap would give a different index of growth. What we would want to do here is to obtain a set of measures which are adequately representative of the many facets of physical development, but which do not overlap to any great extent. Or we could determine what independent factors of physical development we are measuring so that we could compare the children on these factors without having any one of them overly influence our comparison.

Our problem in devising measurement scales in mathematical achievement is similar to this. Devising test items to sample behavior in line with the many goals or purposes of education is like thinking of all the various ways in which physical development can be manifest and which are pertinent to our interests with regard to the physical growth of our children. But many of these behavioral items will overlap to a considerable degree. Therefore, it is necessary to resolve these many aspects into a fewer number of relatively independent components of mathematical achievement.

Test items differ in certain respects, however, from variables or measures such as height and weight. One important aspect in which they differ is that they yield only a two-value discrimination—pass or fail. In order to obtain a scale with greater discriminating power it is necessary to combine scores on several items. Then the number of items passed has the property of an ordinal scale if it can be assumed that success on each of the test items depends in part on a single ability and the amount of that ability needed to pass the item varies from item to item. The extent to which items in a test depend on a common ability is called the "homogeneity" of the test. This property is closely related to, although not synonomous with, that of "reliability". The more homogeneous the test the more reliable will be the scores obtained on the test.

The tasks facing the test constructor are: to take these various items or separate measures of behavior and sort them out so that items which measure the same or very closely related abilities are grouped together to form homogeneous scales; to choose scales which will result in maximum coverage of the domain of behavior in which he is interested and minimum overlap, or else to determine and identify independent components which can be weighted according to their importance.

Two major techniques are available for this task, each of which has some advantages and some disadvantages. One technique is a type of cluster analysis and the other is factor analysis.

The cluster analysis technique makes no __a priori__ assumptions as to which items should be combined to form a test, but attempts to form relatively independent homogeneous scales empirically starting with the matrix of co-variances of the original pool of items. The technique was suggested by Loevinger, Gleser and DuBois in 1953 (2,3). Essentially, it involves building up subtests which have maximum homogeneity and minimum intercorrelations by starting with several nuclei of three items each selected on the basis of having the highest co-variances *inter se* and the lowest co-variances between items of different nuclei of all the available items. Items are added to the nuclei one at a time such that each item which is included adds maximally to the homogeneity of a subtest consisting of the previously selected items plus the one being considered. Excessive overlap of the item clusters is avoided by adding a single item to each nucleus before adding a second item to any one and by excluding items from further consideration if they have already been included in one of the other subtests or if they would fail to make a contribution to the homogeneity of one of the subtests if added at this stage. When all the items have been selected which will increase the homogeneity of these subtests, the remaining items are regrouped to determine whether one or more additional subtests could be formed with sufficiently high homogeneity to merit their use. After all the subtests are formed their intercorrelations are computed and examined for high values. If any are found to correlate too highly, the items forming these subtests are reclustered taking a different pair of nuclei, or else only one subtest is built from the pool. Further improvement in the subtests is obtained by correlating each item in the original pool with each of the subtest scores. Each item in a test should have its highest correlation with that test. An item correlating about equally with two or more test scores is best omitted entirely, since it tends to raise the correlation between the tests.

The clustering method has the advantage that it is completely empirical and also that it results in assigning items directly to subtests which can be used as meaningful factor scores without further statistical manipulation. The reliability of these factor scores can be estimated from the information available in the co-variance matrix. Disadvantages of the method are that the number of subtests obtainable is indeterminate, considerable cycling is necessary to perfect item selection, and a great deal of hand labor is involved.

The second method available for obtaining meaningful achievement scales is factor analysis. Factor analysis is a general name for several techniques which purport to describe the domain of a group of tests in terms of independent or meaningful factors or components and to determine the proportion of variance of each test which can be attributed to each of these factors. From this information, factor scores can be computed from the individual's test scores. The basic assumption is that each measure can be decomposed into a weighted sum of factors common to several of the measures plus a unique factor. This assumption makes the method more appropriate for test than for item scores. In order to use factor analysis in constructing achievement tests, it is best to start with some *a priori* grouping of items into tests. The grouping can be accomplished on the basis of the educational objectives so that all items purporting to sample behavior appropriate to a single objective are scored together. Thus subtests might be constructed for basic skills, quantitative judgment, solving practical problems, transfer of training to somewhat more complex problems, etc. Correlations between each pair of subtests are computed yielding a matrix of intercorrelations. It is this matrix which is factored by one or another technique.

Several different factoring procedures are available. A principle component analysis results in describing a set of n tests in terms of n orthogonal factors, such that as each factor is selected it accounts for a maximum of the remaining test variance. Both common and unique factors are included. The multiple factor method developed by Thurstone (4) attempts to determine only those factors which are represented by at least two of the tests, the so-called common factor space. The latter method permits identification of factors which may themselves be somewhat intercorrelated but which are determined by clusters of tests and which can be meaningfully interpreted.

The choice of the appropriate factor model for a particular problem demands a clear understanding of the purpose for determining factors and the difficulties inherent in each. The ability to discriminate accurately among individual scores on a factor depends on the amount of information about that factor obtainable from the original tests. Factors which are represented in several tests generally yield more reliable scores than those which are represented by at most one variable. Thus, while a principle component analysis results in factor scores which cover the entire domain that is being investigated, these scores differ considerably in their reliability and some factors may yield almost no reliable discrimination. On the other hand, multiple factor analysis, while possibly resulting in more reliably determined factor scores can result in the omission of important dimensions of the original domain if these dimensions are each represented in only one test. Furthermore, decisions as to the number of common factors to extract is difficult to make in practise. The final factor solution is not unique, but must be obtained by rotation of axes. No simple single criterion can be used to determine the optimum resolution in every case. Subjective considerations such as meaningfulness and interpretability enter into the final choice.

In conclusion it should be noted that both cluster analysis and factor analysis are based on the correlations among items or groups of items for a sample of individuals. The results obtained depend to a great extent on the teaching methods employed with the individuals in the sample. Any particular teaching method, by virtue of approach, emphasis, and order of presentation, tends to determine what responses develop together. Intercorrelations among items for students in any one class will likely reveal a single predominent factor reflecting general learning ability. Thus, in choosing a smaple for the development of achievement scales by these methods it is important that consideration be given to the adequacy of representation of different classes and different teaching methods.

While both cluster analysis and factor analysis present some technical problems, they both offer possibilities for substantially improving our measurements of mathematical achievement. Such techniques, coupled with a thorough study of the desired objectives of a mathematics curriculum, careful construction of items, and an adequate and representative sample of students from many different classes reflecting different teaching methods upon which to base the necessary statistics, should result in more precise scales with which to evaluate educational research in the field of mathematics.

REFERENCES

1. Brownell, William A. and Moser, Harold E. Meaningful versus mechanical learning: a study in grade III subtraction, Duke University Research Studies on Education #8, 1949.

2. DuBois, Philip, Loevinger, Jane, Gleser, Goldine. The construction of homogeneous keys for a biographical inventory. Human Resources Research Center. Res. Bull. 52-18, May, 1952.

3. Loevinger, Jane, Gleser, Goldine, DuBois, Philip. Maximizing the discriminating power of a multiple-score test. Psychometrika Vol. 8, #4, Dec. 1953.

4. Thurstone, L. L. *Multiple-factor Analysis* (Chicago: University of Chicago Press, 1947).

MENTAL ABILITIES AND PERSONALITY VARIABLES IN RELATION TO PROFICIENCY IN MATHEMATICS, WITH PARTICULAR REFERENCE TO THE COLLEGE LEVEL

S. B. Sells
Texas Christian University

Introduction

In connection with the supervision of a thesis research project by a graduate student (Louise G. Yates) the writer has during the past year extensively explored the general area of prediction of success in college level mathematics. The selection of this level as the focus of the investigation was an expression of our interest, but we were surprised to find that very little competent work has been published in this area.

The broad rationale of our approach was as follows. Since capacity and aptitude are generally accepted as significant parameters of learning, they must be considered in the analysis and evaluation of mathematics training. In addition to the identification of the ability and personality factors which predict mathematics proficiency, the approach is relevant to student personnel selection, curriculum organization and methods of teaching mathematics.

As in the case of most prediction problems, prediction of mathematics proficiency appeared at first to be a simple and straightforward problem; that is, until an effort was made to define the construct and to identify appropriate criteria to measure it. This paper will attempt to show that the definition and criterion problems are major issues which merit thoughtful study before any analyses of relevant predictors are undertaken.

The first source consulted concerning the definition of the universe of content representing mathematics proficiency was a small sample of university mathematics professors. Somehow we had naively assumed that the venerable "mother of sciences" was a well structured, firmly established, solid core, content discipline and that its solid core could be specified by any of its qualified practitioners. It was therefore something of a surprise that there was as much variation among mathematicians' opinions as one usually encounters among those of psychologists.

At least one vote each was cast for including plane, solid and analytical geometry, algebra, trigonometry, calculus and advanced calculus as components of the general content universe. However, the mathematicians disagreed widely on the total definition. The informal interviews also disclosed variation in the mathematical processes included under various course titles. And finally, further variations of opinion were voiced on questions of mathematical proficiency essential to a number of applied fields, such as engineering, statistics, operations research, and sciences, such as physics and astronomy.

Fortified with the reassurance that mathematicians were divided on the questions raised, illumination was next sought from the research literature.

The results of an extensive search are summarized briefly.

(1) Attempts to define a construct of a universe of mathematics performance have been concerned more with mathematics aptitude than with mathematics proficiency. Concepts based on reasoning ability have been most frequently cited. Development of a predictive criterion requires emphasis on performance rather than aptitude.

(2) Three types of criterion measures have been used in most published studies (see selected references). The most common consists of course grades in mathematics. Ratings of students by instructors and achievement tests have been used occasionally. The limitations of grades need no discussion here. No criteria have been reported, however, which represented even an attempt to define the universe. The suspicion seems justified that what has been predicted can only vaguely be described.

(3) No systematic analyses of quantitative behavior, either at the earlier levels or at the college level were found. There is a great need for such research, to provide satisfactory instruments for the measurement of learning progress and performance in mathematics in various contexts.

(4) Results of extant studies, with the limitations mentioned, indicate that V (verbal comprehension) and R (reasoning) abilities and study skills, possibly reflecting motivational patterns, have generally reached substantively and statistically significant predictive levels. N (number facility) has been found relatively unimportant beyond simple elementary arithmetic. Some personality factors have reached significant levels, but no stable patterns have yet been reported.

Definition of Mathematical Proficiency

Our review of literature, experience with our own research project and reflections on these have led to the exploration of a number of questions relevant to research on the definition of the criterion universe mentioned above.

First, we have been impressed with the extensiveness of forgetting, even among mathematics majors. The achievement on an analytical geometry test by students who had completed the course a year previously was impressively poor in a substantial number of cases. When forgetting is high and also widespread, one may well question both the students' degree of mastery of required content of a course and the extent to which such mastery is either an adequate prerequisite to later courses in a sequence or transferred to other tasks.

This observation recalls Skinner's discussion of grading practices in his article on "Teaching Machines" (Science, 24 October 1958). If learning progress were regulated by the requirement of complete mastery of each earlier phase of a subject before the student were admitted to the next phase, in contrast to the traditional school system of advancement by achieving a passing grade (i.e. exceeding a minimum standard), control of learning progress and measurement would be more rational. If a student presently progresses with B and C grades, his training may be presumed to be quite incomplete, even without forgetting.

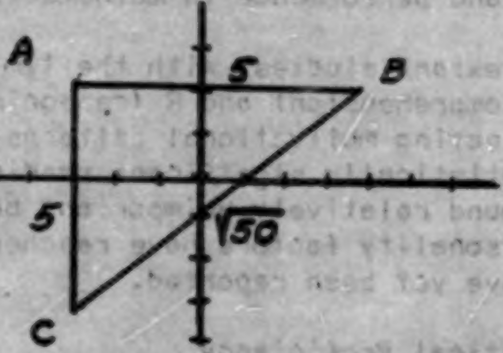
Emphasis on completion of units would have the effect, also implied when attempts to define teaching programs are made, of requiring careful reevaluation of course objectives and content. The ordering of content sequentially is of course one of the most important issues. The investigation of the reorganization of the classroom along lines suggested here is considered a promising research problem.

Second, we have had occasion to question the meaning of mathematics proficiency as reflected in conventional achievement tests. For example, the following item appeared on an achievement test for analytical geometry:

If A = -4, 3; B = 3, 2; and C = -2, -3;

What is the length of BC?

This question can be answered directly in analytical geometry by a distance formula. It can also be answered by plane geometry by plotting the points with reference to rectangular coordinates, as shown below, and computing the value of the hypotenuse of a right triangle.



If the examinee were unfamiliar with the conventions of analytical geometry he might be unable to solve the problem as given above. On the other hand, if he were presented the same problem with the coordinates as shown, he might solve it correctly even if he had forgotten (or not learned) the distance formula.

This example demonstrates the need for distinguishing between formal procedural knowledge, on one hand, and quantitative behavior on the other. The following implications of this distinction are relevant: (1) different solutions of the same problem are classified under different branches of mathematics; (2) convention or congruity with other procedures may dictate the particular approach that is taught or recommended in various fields which use mathematics as a tool; (3) since mathematics is important in the training of scientists, engineers and other workers in practical fields, apart from the training of mathematicians, the conventional mathematical classifications and formulations may be inappropriate concepts for defining a universe of "quantitative behavior."

There appears to be a discrepancy between a number of manifestations of numerical facility, quantitative thinking, and quantitative problem solving, on the one hand, and both attitude and proficiency in formal mathematical operations, on the other. In attempting to think this through the following questions were posed. (a) To what extent are the anecdotal observations valid which suggest that many individuals perform various particular quantitative operations (e.g. making change, balancing accounts, computing airspeed and navigating, etc.) successfully who either dislike or are unable to succeed in formal mathematics? (b) Assuming their validity for the sake of discussion, what hypotheses might account for them? (c) What implications do these have for the issues discussed here? These are discussed below.

(a) A survey of a number of age, sex and occupational samples might be valuable in opening up this line of approach and also in obtaining more accurate data on the nature and extent of the suspected discrepancies. It would be desirable in data collection to preserve the situational contexts of the various activities sampled, e.g. making change with money, balancing actual own check books and accounts, etc.

(b) It is suspected that in formal mathematics quantitative proficiency is an end in itself whereas it may be subordinated to a situational problem in other fields. To the extent that this is true, the first hypothesis is that the situational context for mathematical operations may be an important variable, highlighting differences in "need to know", range of mathematical operations called on, and degree of overlearning of specific techniques.

A second hypothesis relates to the use of mathematical symbols in mathematics courses. These represent a high level of abstraction in contrast to more concrete and "semantically slanted" formulations in applied situations. Related to the use of mathematical abstractions is the oft-noted tendency of mathematicians to elude and to omit "obvious" intermediate steps in exposition of derivations or proofs and in teaching. This may create problems of understanding their communications and possibly contribute to what has been called "number anxiety". The abstract simplicity of symbolic notation is often a difficulty to learners. Attention needs to be drawn to the distinction between simple in the sense of economy of explanatory dimensions and simple as a synonym of easy.

A third hypothesis concerns the dependence of more advanced mathematics on basic operations. It is known that forgetting of school curriculum content is rapid where the material is not used and is displaced by currently more functional material. This hypothesis assumes that much of the mathematics content is forgotten and that such forgetting creates problems in later learning requiring mastery of basic material.

A fourth hypothesis concerns the use of computing aids that are specifically tailored to particular problems. For instance, the bookkeeping or billing clerk does not use a desk calculator, but a billing or accounting machine; the pilot does not use a slide rule, but an air-speed computer. The hypothesis is that such tailored computing aids may be learned and used proficiently in the context of a particular purposive situation more readily than general computing devices based on general mathematical principles.

The third general question in this section concerns issues involving motivation. Some writers have mentioned the reputedly widespread fear of mathematics among teachers as well as students and in the population generally. As suggested above this rests on anecdotal evidence which apparently refers to formal mathematics primarily. Many individuals who report such fears or dislikes of mathematics regularly perform quantitative tasks in their work, sports, business affairs and recreation.

Proficiency Criteria and Prediction

For prediction of success in mathematics in both pure and applied references, using ability, personality and interest measures as predictors, definitions of proficiency in quantitative behavior are essential to the development of criteria.

The research approach which seems most clearly indicated to meet the apparent need is first, the development of a large pool of test items, representing diverse variations of selected core quantitative problems, with reference to amount of information given, problem and language (situational) context, and alternative methods of solution possible. These should be administered to different samples of students with varying training by major subject, with appropriate control of general mental ability, background and training, and time elapsed since completion of relevant courses. Multivariate analysis of item covariances in parallel samples may point the way to the ordering of results into meaningful groupings of items which could then be scaled and standardized as proficiency tests. Although available performance criteria such as grades and ratings of professional competency are fallible, correlations of tests with such measures would be of interest for secondary external validation.

With suitable proficiency tests, factored ability, personality and interest tests could be employed in routine prediction studies and hypotheses suggested by present literature and further research could be tested.

SELECTED REFERENCES

Barakat, M. K. A Factorial Study of Mathematical Abilities. Brit. J. Psychol., Stat. Sect., 1951, 4, 137-156.

Barrett, Dorothy M. Differential Value of Q and L Scores on the ACE Psychological Examination for Predicting Achievement in College Mathematics. J. Psychol., 1952, 33, 205-207.

Bromley, Ann and Carter, Gerald C. Predictability of Success in Mathematics. J. Educ. Res., 1950, 44, 148-150.

Coleman, Ralph H. An Analysis of Certain Components of Mathematical Ability and an Attempt to Predict Mathematical Achievement in a Specific Situation. Dissertation Abstracts, 1956, 16, 2062.

Curry, John Foster. The Effect of Reading Instruction on Achievement in Seventh-Grade Arithmetic. Dissertation Abstracts, 1955, 15, 2059.

Hills, John R. The Relation Between Certain Factor-Analyzed Abilities and Success in College Mathematics. Univ. S. Calif., Psychol. Lab. Rep., No. 15, August, 1955.

Kinzer, John R. and Kinzer, Lydia G. Predicting Grades in Advanced College Mathematics. J. Appl. Psychol., 1953, 37, 182-184.

Krathwohl, William C. Relative Contributions of Aptitude and Work Habits to Achievement in College Mathematics, J. Educ. Psychol., 1953, 44, 140-148.

Lee, Doris. M.A. Study of Specific Ability and Attainment in Mathematics. Brit. J. Educ. Psychol., 1955, 25, 178-189.

Murray, John E. An Analysis of Geometric Ability. J. Educ. Psychol., 1949, 40, 118-124.

Oickers, P. J. A Factorial Study of Arithmetical Ability. J. for Soc. Res., Pretoria, 1951, 2, 1-21.

Pitts, Raymond J. Relationships Between Functional Competence in Mathematics and Reading Grade Levels, Mental Ability, and Age. J. Educ. Psychol., 1952, 43, 486-492.

THE INTERACTION OF PERSONALITY VARIABLES
WITH TEACHING METHODS AFFECTING
ON COLLEGE LEARNING

W. J. McKeachie
University of Michigan

For some time I have been studying methods of teaching psychology. A series of experiments convinced me that research on teaching methods per se was a rather unrewarding activity. I was luckier than many investigators in getting some positive results, but nothing I found was world-shaking. As Dr. Carpenter has pointed out, most studies of teaching methods have been unable to reject the null hypothesis. Whether the problem is large vs. small classes, lecture vs. discussion, student-centered vs. instructor-centered teaching, television vs. live classes, independent study vs. regular classes, the results almost invariably are "No significant differences!" Dr. Carpenter's suggestion is that we need to develop larger conglomerates of variables if we are to show a significant effect, and there is much to be said for this approach. But there are other factors which may also have contributed to negative results.

Experienced faculty members and administrators have known for years that no one teaching method succeeds with all kinds of students and that different teachers are not equally effective with all methods of teaching. There is some research support for these beliefs. For example, Wispe (1951) found that anxious students dislike discussion teaching, and Smith's (1956) results support this. Perhaps instead of asking "Can we teach as well by Method A as by Method B?", we should ask such questions as, "With which method of instruction is a particular teacher most effective?" "Which method of instruction is best for student X?"

Moreover we might expect that teaching methods effective in a course with one type of content might be ineffective in a course with another type of content. Even within a course we might expect that methods appropriate for one type of objective, such as knowledge of rules, might be inappropriate for other objectives, such as application of rules. Thus I propose that the significant effects of teaching are interaction effects between student characteristics, teaching method, teacher personality, and task.

This makes research in teaching methods enormously complex, but there are still at least two other barriers to effective research on teaching. One is definition of the independent variable, teaching method. All too often we compare two methods which are only vaguely defined, and two studies come out with apparently inconsistent results because experimenters used the same title for methods which were quite different. Moreover, the methods chosen for comparison sometimes have little relevance to any psychological or educational theories which would permit generalization beyond the experimental situation. This is one of the things we should keep in mind if we follow Carpenter's suggestion.

The second barrier is the lack of adequate measures of educational outcomes. Typically we use as our criterion measure the final examination in the course. Not only is this measure likely to test only the lowest levels of achievement, but it is also a very insensitive measure of experimental effects. Let me explain why I say "insensitive."

Whatever a student's motivation for being in college, grades are important. If he is really interested in learning, grades represent an expert's appraisal of his success; if he is interested in getting into professional school, good grades are the key which will unlock graduate doors; if he wants to play football, grades are necessary for maintaining eligibility. In short, most students are motivated to get passing grades.

Because grades are important to students, they will learn whatever is necessary to get the grades they desire. Their best clues are examinations. If we base our examinations on memorization of details in the textbook, students will memorize the text. If they believe our grades are based upon their ability to integrate and apply principles, they will attempt to do this. If we teach poorly, students will simply study harder. As long as students know what the final exam will cover, they will devise legal and illegal ways of doing their best on it.

Thus the final examination is not only not designed to measure the most important outcomes but also is so greatly affected by the student's motivation for a good grade that it's not a sensitive measure of most teaching method differences. In fact I suspect that when we do find differences in final exam scores, it is often because one class simply knows better than the other what to do to pass the examination.

Last year support by the Ford Fund for Advancement of Education enabled us to investigate the interaction between student personality and teaching methods as they affected learning in three multi-section courses. We felt that students' needs for achievement, affiliation, and power and student anxiety about failure were among the most important personality characteristics influencing learning. We then looked for teaching methods which would provide cues for these motives, hypothesizing that an instructor who made frequent reference to standards of achievement would elicit achievement motivation; one who was friendly would elicit affiliation motivation; one who encouraged student assertions would elicit power motivation, etc. We chose three courses differing in goals and content because we presumed that a high level of anxiety might facilitate relatively rote learning while disturbing more complex learning.

Our project involved observation of 30 instructors teaching in 3 large courses: elementary psychology, college algebra, and second year French. Pairs of observers visited each class. One observer used Bales' technique of counting teacher and student acts falling into categories such as "expressing agreement," "giving suggestions," "giving information or opinion," etc. The other observer noted whether the instructor followed an outline, called the students by name, etc. At the end of the course, we also got student perceptions of the same behaviors. The scale used by the students was developed by factor analysis from a longer scale listing behaviors we thought to be relevant to whether or not the instructor would arouse achievement motivation, affiliation motivation, or anxiety about achievement.

To measure student personality we used the Thematic Apperception Test of need for achievement, need for affiliation and need for power developed by McClelland, Atkinson, et al. In addition we used the Taylor test of manifest anxiety, the Haber-Alpert anxiety scales, and the California Psychological Inventory. Finally we interviewed each instructor to determine his goals and his perception of his class. Thus our design involved four types of independent variables - student personality, teaching methods, teacher characteristics, and task - and three types of dependent variables - achievement on final examinations, student satisfaction, and attitudes of students towards the content area (e.g., psychology or mathematics).

The results of our interaction analyses are not yet completed, but the first one came out with negative results. Consequently I am not able to give you any hot-off-the-IBM 650 information about what method to use with what students. However, I can tell you the characteristics of the mathematics teachers who were judged by students to be most effective.

Characteristics of Mathematics Instructors Rated as Being Most Effective.

Students rate the mathematics teachers they like best as providing more structure and organization than the poorer instructors. The observation records of the highly rated teachers showed little inattention, a high degree of student participation, and a good deal of informal encouragement. These, incidentally, are characteristics on which most mathematics classes are low as compared with the other courses we observed.

We have some evidence from other courses that student ratings of effectiveness have validity. We had tried to get each course to differentiate between low level and high level outcome measures. In French we had oral and written examinations; in psychology, an objective test and an essay. We then ranked teachers for effectiveness in terms of mean student performance on the tests, partialling out intelligence. In French, the most effective teachers as measured by student performance on one examination were also those most effective as measured by student performance on the other. Student satisfaction with structure and standards correlated significantly with effectiveness.

In psychology, however, our results were more puzzling. Those instructors who were most effective as measured by their students' performance on the objective test tended to be the least effective as measured by their students' performance on the essay test. Apparently in talking about teaching effectiveness we need to ask, "effective for what?"

Sex Differences in Motivation and Performance.

In view of Dr. Taylor's discussion of sex differences, I would like to add some odd findings of sex differences in our studies. Richard Carney, one of our graduate students who is now at Drake University, has been studying motivational factors affecting achievement. He compared performance on "practice" tests with performance on "credit" tests.

He hypothesized that credit tests engage so many motives that no single motivational measure will predict performance. Because most students are highly motivated, they perform close to capacity, and intelligence thus should correlate with performance. On practice tests, however, fewer motives should be engaged so that measures of single motives may predict performance. In so far as the motives engaged are uncorrelated with intelligence we would expect the predictive value of intelligence tests to be less than in the credit situation.

Carney's results were neatly confirmed for men. For women, however, intelligence tests predicted well in both "credit" and "practice" situations. Apparently his women students were working close to capacity in both situations. This might be accounted for by the hypothesis that women are so anxious about achievement that they do not discriminate as clearly as men between different types of achievement situations.

Dr. Taylor has presented data indicating that more masculine women are more effective problem solvers than more feminine women. We have some which may indicate some qualifications of his conclusions.

Robert Abelson, one of the graduate student assistants in our undergraduate psychology statistics class, was concerned about the blocks of which girls taking statistics complained. He felt that such blocks might be the result of conflict over assuming the feminine role, and that as compared with other women, girls with such conflicts might perform more poorly on a test stated in mathematical terms than on a test stated in terms of familiar housewife-type problems. (I predicted just the opposite.) He thus developed parallel forms of two tests. One form of each test stated a question about buying yard goods or groceries; the other stated the problem abstractly. Each girl took one of the tests in mathematical form; the other in the housewife form. Rejection of the feminine role was scored (blind) from answers to open-ended questions. The results gave clear-cut support to his hypothesis. "Non-housewife" women did relatively poorly on the math form. Moreover in another group Abelson found that the women who rejected the housewife role had lower levels of aspiration for mathematics achievement than did those who accepted the role.

Although we probably cannot do our mathematics teaching all in terms of buying groceries or draperies, Abelson's experiment does furnish a dramatic illustration of my main point: individual differences in personality make a difference in student learning.

REFERENCES

Smith, Donald E. P. Reading Improvement as a Function of Student Personality and Teaching Method. The J. educ. Psychol., Jan., 1956, Vol. 47, No. 1, pp. 47-59.

Wispe, L. G., Evaluating section teaching methods in the introductory course. J. educ. Res., 1951, 45, 161-186.

PERCEPTUAL FACTORS IN THE LEARNING OF MATHEMATICS

James M. Vanderplas
Washington University

When I was asked to present a paper to this group on perceptual factors in learning mathematics, I agreed to do so with some misgivings concerning the outcome. What could a psychologist say, about the relation of mathematics to perception, that had not been said many times before, and with greater force and clarity? Mathematics, as a formal discipline, is far older and more mature than is the scientific psychology of perception; the methods of proof are relatively rigorous and well-defined; and the teaching of mathematics would appear to involve primarily the communication of established principles from teacher to student. The scientific psychology of perception, on the other hand, is adolescent by comparison; although its methods are sometimes rigorous, its concepts are not always well-defined, and its teaching often involves the grasping of concepts by intuitive understanding rather than clearcut, formal communication.

Theories of perception are heterogeneous and often in disagreement concerning the invariant relations underlying common and obvious facts. Perception has been dealt with in psychological theory at various levels, ranging from simple sensory aspects to those involving the roles of complex value systems, individual needs, and heterogeneous motives. Allport (1955) recently discussed in detail thirteen separate and distinct theories of perception and found it necessary to construct additional concepts to encompass the range of phenomena embraced by these theories. Bartley (1958) chose to add to these a fifteenth, perhaps more all-embracing, theory of perception.

In each of these theories, one finds limitations: An approach which considers only the sensory aspects of perception ignores much of the influence of past experience on present perception, while a directive-state theory, emphasizing needs, values, and motives, clearly neglects many obvious factors related to the organized character of the relations among parts of the stimulus-object.

In bringing to bear the findings of such a segmented field upon problems of learning mathematics, we should also keep clearly in mind that the factors of perception are often little more than isolated demonstrations, descriptions of phenomena, or selected findings deduced from hypotheses of only limited generality.

In discussing the role of perception in learning mathematics one also encounters some difficulty in definition. In its broadest sense, one might argue that mathematics, after all, _is_ perception. For it is _with_ mathematics that the statement of invariant relations among variegated, concrete objects, is made; and it is this kind of statement that makes meaningful a world where change and heterogeneity are the rule. Out of this process of statement of invariance common features, such as number, or form, or mass, are established; and the abstractions of these invariants, in the symbolic language of mathematics, become the elements of the concepts employed in mathematical reasoning. Perception is thus understood with mathematics, or, as Dr. Edgar Anderson so aptly expressed it, perception is to be understood when treated as mathematics.

In another broad sense, once these abstractions arising our of mathematical formulations are stated symbolically, mathematics becomes divorced from so-called "reality" and bears little, if any, relation to the world as perceived via our senses. What, for example, is the sensible counterpart of a point? or infinity? or $\sqrt{-1}$?

Perception itself is not an easy concept to define; definitions range from those which consider it as immediate awareness or cognition of relations among sensations to those which regard as relevant only responses, clearly observed, which correlate with specifiable change in stimulus-object relations. Between these extremes lie definitions which characterize perception as an intervening process, variously described, which mediates adjustive acts in the presence of objects or on the occasion of events. The latter kind of definition seems to fit well into the schema of our present topic; for a process, thus characterized, would include many aspects, ranging from simple sensory relations corresponding to stimulus dimensions such as size or direction, which mediate adjustment to physical space, to complex behavioral relations which depend on culturally-derived experiences and which mediate adjustment to social stimuli such as the behavior of other persons.

We shall, of course, be selective, focussing principally upon aspects relevant to mediation of adjustive acts in learning mathematics. We shall therefore be considering some selected factors of perception, related to objects and events which influence the attainment of concepts by the learner.

Looking at mathematics briefly from an historical point of view, one may find many relations of this sort. The development of mathematics may be regarded as based upon the relation of its symbols to the way in which man perceived the world about him. The first mathematician could have been the genius who conceived the artful device of substituting a mark on a stone for the pebble in the bag which was his computer used to count his sheep (Smith and Ginsburg, 1956). As late as 1520 or so, no course in numbers was complete without an account of the process of counting on one's fingers. Many of the words used in denoting the abstract concepts of numbers also denote objects of the perceived world. The Russian word piat (five) resembles piast, or outstretched hand, and it is easy to duplicate many roman numerals by arranging the fingers, hands, and arms appropriately. Our number system thus appears intimately related to perception of objects; one historian speculated that we might have evolved a different system if we had been endowed with six fingers (Dantzig, 1957), and one can even consider the possibility that the invention of shoes prevented the development of a system based on 20 rather than 10.

One may also speculate on the historical development of geometric concepts: Euclid, for example, may have defined a point and a line to represent real objects such as a dot or string. He may have transformed the definition into an abstraction only after its logical difficulties were discovered.

We are, of course, speculating. Helmholtz, in his masterful analysis of the development of many of the axioms of geometry, stated the assumption more aptly. He said, "...we are stopped in our power of representation, by the structure of our organs and the experiences got through them which correspond to the space in which we live." (Helmholtz, 1956, p. 655.)

Thus, if our assumption is correct, we see that despite the (largely) abstract character of much of mathematics many of its concepts are rooted firmly in sensory-perceptual processes and are tied closely to the world of objects and events. It is those objects and events, principally those which are or which may be employed as aids to teach mathematics, and their relation to perception and the attainment of concepts, to which I would like to turn. Before doing so, however, I would like to say a few words about the recent growth and emphasis in the field of mathematics of the use of sensory-perceptual aids.

There was a time, not too long ago, when the only aid to teaching mathematics was a textbook dealing solely with the formal abstractions. Teachers today make use of many visual aids, learning guides, and manipulative tools to augment formal instruction with demonstrations and student participation. The use of these materials, as stated by modern textbook writers (e.g., Brueckner, et al., 1957), is designed to foster appreciation of the uses of mathematics in everyday life, vitalize instruction, make mathematics realistic and meaningful, and aid in the development of conceptual understanding. Learning, through direct contact with objects and events related to the experience of the student, is thus integrated with vicarious learning through study of books and printed materials.

So great has been this emphasis on teaching aids that one can find some form of sensory-perceptual aid suggested for almost any kind of mathematical concept. The Commission on Post-War Planning of the National Council of Teachers of Mathematics found in 1945 that well over 2,200 filmstrip subjects had been produced by the Navy alone; they stated that the figures for visual aids, charts, graphs, models and other devices would exceed the imagination (Commission on Post-War Plans, 1945). Continuing emphasis on use of these materials is evidenced in several monthly departments on teaching aids which appear in the mathematical periodicals (e.g., Syer and Johnson, 1948; Berger, 1950).

Part of the stimulus to this growth in recent years may be due to the devotion of a full volume, on the use of sensory aids, of the Yearbook of the National Council of Teachers of Mathematics in 1945. Many of the papers contributed to that volume urged the use of sensory aids and suggested examples of such aids.

Additional stimulation to the growth in the number of sensory-perceptual aids may be traced to the corresponding growth of information in the field of perception. Theories of perception have become more analytical in recent times and have dealt with situations and relations more likely to be of a sort which yields quantitative data applicable to the design of teaching aids. An example of this trend is the recent experiment by Wallach and O'Connell (1953) in which there was isolated the effective stimulus for a report of three-dimensional appearance of lines and geometric figures in a dynamic two-dimensional display. The effective stimulus reported by these investigators was the simultaneous change in both angular position and length of the lines as a function of time. When either was absent subjects reported merely a distorting two-dimensional object. The work of J. J. Gibson (1950) in delineating some of the more complex relations of motion, gradients in spatially organized events, over time, is also representative of this trend. It seems likely that these relations would influence the effectiveness of many aids for teaching (say) geometry; the data of these experiments could be used in the design of aids such as motion pictures.

Other studies, such as those by Bruner and his associates (e.g., Bruner and Goodman, 1947), have yielded increasingly analytical information about the role of past experience in perception. One study, though not conclusive, found that "poor" children tended to overestimate the sizes of coins more than "rich" children, the more so the greater the denomination of the coin. Such a finding, applied to the design of a teaching aid, would indicate that one should select rather neutral objects as examples of a concept of size, related to number, to reduce the possibility of adverse influence of such social factors.

In general, then, it would seem that the trend is clearly in the direction of increased use of sensory-perceptual aids, of greater effectiveness corresponding to increased knowledge about the processes of perception. It would also seem that, as we suggested earlier, little might be said beyond urging continued and extended use of such aids. As more information becomes available and theories of perception are clarified and new principles develop, one may expect the needs of the teacher to be increasingly satisfied by application of the data and principles. Enthusiastic and increased use of such applications would seem to provide an obvious answer to many of the problems in learning mathematics.

But suppose that one concludes such advice on this note. Is this to say that the teacher has now only to look for a perceptual aid and he will find one? Well, yes and no. For having looked for and found an aid for teaching the concept in question, he will undoubtedly have found as well, argument for use of several others. Also, he will in all likelihood have found an argument against the use of any concrete model or aid to represent the concept, perhaps an argument for the use of a process, to facilitate the discovery of the concept, with no apparent aid at all. Further, it is likely that he will have found conflicting evidence for the effectiveness of several similar aids in attainment of the concept. How does he choose among them? I hope that I do not parrot the remarks of other conferees when I say, "I don't know." I do know, however, that I cannot, at this point, make a general statement concerning the usefulness of perceptual aids without a detailed examination of their characteristics and the relation of these to the particular concept in question. What we are emphasizing, of course, is the obvious point, often overlooked, that some aids facilitate attainment of some concepts some of the time, for some people, while others facilitate others.

I would like, at the risk of laboring the obvious, to cite an example in the form of a recent experiment: In the experiment the ability of kindergarten children to recognize the number of objects in a picture and to draw a corresponding number in a second picture was studied as a function of the kind of objects and the context in which the objects appeared.

It is fairly well agreed that many children can perform these tasks without counting, treating the objects as a group, while others perform the task by counting, treating the objects one at a time, in serial fashion (Carper, 1942). The question, then, of the experiment, concerned the effect of the method of presenting the objects upon the development of grouping behavior.

When the objects were simple geometric forms, grouped into a single pattern such as a triangle or square, about 50 percent of the children studied were able to recognize 3-ness or 4-ness without counting. When the same groups of figures were shown in context, as parts of an object, such as a tree or a savings bank, about 38 percent of the children were able to recognize the number. When the objects were further imbedded in a more complex context by adding other objects, the percentage of children who successfully grouped dropped to 29. When the objects were further elaborated by representing them as people or animals instead of geometric forms and imbedding them in a picture such as a farm scene, the percentage dropped further, to about eleven.

Both the kind of object used as an example of the concept (3-ness, 4-ness, or 5-ness) as well as the context in which it appeared, may be seen to impede rather than facilitate the attainment of the concept. The specific relation, between the example and the concept, would thus appear to be an important consideration in the selection of materials for use as teaching aids. A concept in mathematics may be abstract while the model employed to illustrate it may be concrete and related to the generalized concept in only a limited way.

Some abstractions may be illustrated well by a concrete example, others not. Consider as an example the concept of a closed, continuous line of infinite length which bounds a finite area. I have seen an illustration of this concept which makes use of a series of drawings. Beginning with a triangle, the illustration proceeds through a series of modifications of the form in which increasingly many, smaller and smaller, indentations and projections are made on the perimeter, thus increasing the apparent "jaggedness" of the figure while holding the area constant (Syer, 1953). Although this particular figure seems to me to represent the development of the concept very well, I have no assurance that it will be equally effective for all persons nor that similar figures will be effective for other similar concepts, without an evaluation of the figure.

A recent experience of mine reminded me of the need to consider carefully the relation between the figure chosen as an illustration and the concept illustrated. I had just finished explaining to my statistics class that the joint function of two independent frequency functions could be schematized as a surface in three-dimensional space, the volume under the surface representing a probability in the same way that the area under a curve represents a probability in two-dimensional space. One student, who appeared particularly unimpressed even after I sketched a perspective view of the figure on the board, said to his neighbor, "I think I'd prefer a thousand words." Later questioning brought to light that although the sketch was clear enough, the relation between graph and formula was not. The analogy somehow just didn't fit.

This experience reminded me that sometimes the illustration of a concept does not adequately represent that concept to all persons. An added complication may also arise in some instances, if the illustration represents a concept other than the one in question.

Dr. Edna Heidbreder has conducted an extensive series of experiments on factors related to the attainment of concepts, numerical and otherwise. One in particular illustrates nicely how the concept attained, as a common property of a variety of instances may depend on the character of those instances. She used a pack of 144 cards, each card with a different drawing on it, and instructed the subjects to sort the pack into 9 piles. The piles were to be sorted according to a classification to be worked out by the subject. In the pack were three kinds of concepts: Things, consisting of pictures of 3 kinds of objects, hats, faces, and birds; 3 kinds of nonsense forms, wiggly lines, abstract shapes, and patterns; and 3 kinds of number concepts, 3's, 4's, and 6's. The question was whether the subject could abstract these properties without specific instruction. The subjects, college students, showed 81 percent correct classification of thing-concepts, 39 percent correct for nonsense forms, and 32 percent correct for number concepts (Heidbreder, 1948).

Things, perhaps, lie closest to the world of concrete objects with which we have had most experience and thus are abstracted and used with greater facility. Number, on the other hand, must be divorced from the concrete object, and thus it may represent an intrinsically more difficult concept to attain. It is perhaps for this reason that teachers of mathematics find it helpful to employ models of concrete objects to illustrate many of the concepts of their subject matter. However, we would also suggest that care must be taken in the selection of particular models lest the learner develop the thing-concept and overlook completely the numerical or symbolic abstraction of which the model is an example. Heidbreder's experiment suggests that irrelevant concepts may emerge as a result of use of certain materials as teaching aids. A related point that we may also raise is that an irrelevant perception may result, from the use of a model, even when the model is clearly related to the concept.

A commonly used visual aid in demonstrating a cycloid curve involves the use of a disk with a dot on the rim. The circle is rolled along a line to demonstrate the arc through which the dot moves. The perception of a cycloid movement, however, has been shown to depend upon the conditions under which the demonstration is viewed. In a darkened room, with only the dot visible, the cycloid movement is clearly seen. If another dot is added, at the center of the circle, one might predict that the observer would report two movements: one of the dot in a cycloid path, the other of a dot in a horizontal path. This is not the case. Observers report that the peripheral dot appears to move in a circle around the central one and that the whole "system" seems to move along together. The cycloid movement is absent (Krech and Crutchfield, 1958, p. 125). In an ordinary classroom demonstration, then, one can wonder if the student actually "grasps the idea" of a cycloid from a demonstration of this sort.

Although the evidence for failure of the cycloid to be apparent is not conclusive, further support for the idea that a concrete model may impede perception of the curve comes from the experience of a teacher who found students in heated disagreement about the path of a point on the circumference of an automobile wheel. Some thought it would move in a path parallel to the surface of the road, others thought it would move in a circle, still others believed, wrongly, that it would trace a series of loops, such as those drawn in penmanship classes (Young, 1945).

Along these same lines, we may mention a case or two which illustrate that some teaching aids may exemplify relations which are contrary or unrelated to the operations carried out in connection with the concept. Several textbooks were found by one writer (Fehr, 1953) which used illustrations such as the following to represent the process of subtraction: 000000000 - 000 = 000000. It would seem that a simple demonstration, in which the teacher simply removed some marbles or other objects from a group, would serve much better in this instance.

A second example of this lack of relation comes from the experience of some scientists (cf. Newson and Randolph, 1946) who report difficulty with the equation $Y = \sin X$ when graphed in cartesian coordinates because of their inability to "accept" the device of marking X on a straight axis. Probably all of us who learned the definition of sin X learned it in terms of the ratio of the opposite side to the hypotenuse of a right triangle, despite the fact that there are at least ten other ways of defining it, only one or two of which have any relation to angles (cf. Cell, 1950). To define the trigonometric functions in terms of angles in about as reasonable, according to Newson and Randolph (1946) as "...an attempt to define arithmetic as the science of money....", since only about five percent of the problems in both areas involve these applications.

What we have said thus far would seem to suggest that careful attention should be apid to the characteristics of any teaching aid, preceptual or otherwise, in terms of their relation to the particular concept in question. Some aids are better for development of some concepts, others are worse. Some may hinder the development of the concept or facilitate development of an irrelevancy, and some may foster a false perception or give rise to concepts unrelated to their uses in normal applications. This is not to say, of course, that each and every aid should be subjected to extensive test before it is used; many of the existing aids are so obviously useful that such a critical evaluation is unnecessary. Many of them, further, are in clear accord with accepted principles of perception. The use of colored chalk, for example, or other devices to emphasize figure-ground relations, similarity, and to facilitate "perceptual organization" through contrast and isolation of relevant parts of a figure (e.g., Baravalle, 1944; Blank, 1948; Carper, 1942; Judd, 1932; Lynch, 1942; McCurdy, 1948) are obviously useful devices. Limiting the number of objects in a pattern to (say) six or so, to avoid exceeding normal span of apprehension agrees with accepted perceptual principles, as well as does the use of certain kinds of grouping techniques to increase the number of objects, beyond this limit (Miller, 1956), that can be apprehended.

These principles, however, though well-established, embrace only a limited range of perceptual phenomena applicable to the learning of mathematical concepts. More complex principles, dealing with such factors as past experience and set, less well-defined, allow considerable room for error in judging the effectiveness of a teaching aid. I would, therefore, like to cite an example or two which illustrate the way in which these factors may operate adversely in the development of concepts, to stress the need for evaluation when any but clearcut relations may be observed between the teaching aid and the concept or perception.

My examples bear on the general principle, long-used by teachers, that perception of a new object or event will be influenced by its relation to past experience. This principle is sometimes generalized to imply that any such relation will improve perception and the greater the number the better. Increasing awareness by psychologists of the importance of past experience leads to modification of this principle to include degradation of perception, depending on the type of experience. The study of Bruner and Goodman (1947), already cited indicates that false perception of the size of coins may occur as a result of certain background conditions. The study by Carper (1942), cited above, illustrating adverse effects of surround-complexity upon grouping behavior, also illustrates in part that effects of past experience, with situations depicted in the figures, may be far outweighed by the characteristics of the figure, if not inhibitory in themselves.

Perhaps a better example of these kinds of effects is the following one; it illustrates the operation of a set, based on an appeal to past experience, in hindering the recognition of a clearly defined locus.

A hunter starts from camp and travels 10 miles south. Turning, he travels 10 miles west, whereupon he encounters a bear. Shooting the bear, he transports it 10 miles north, at which point he is back at camp. What color was the bear? If you said, "White," you have established the intended set, and you are ready for phase two: Omit the bear, and require that the hunter travel the same path - 10 miles south, 10 miles west, and 10 miles north. What is the locus of points, on the surface of the earth, which define the possible locations, other than the north pole, of a camp and which satisfy the condition that the hunter may return to it via the stated route? There are, of course, infinitely many of these if this is conceived as a geometric problem. I present this problem to my introductory class in psychology as an illustration of the operation of set in problem-solving. We note, however, that the set is established by reference to a situation which defines (in part, at least) a visual model, by way of our reference to a bear. I doubt that a sketch would be less effective in preventing the student from solving the problem.

A great deal more could be said, both for and against the use of sensory-perceptual aids to teaching mathematics; and I have chosen only a few examples to illustrate the (perhaps) trivial point that not all such aids are as effective as we might want them to be. We would urge continued and increased use of such aids, but we would underline the qualification, where they are clearly effective in developing the particular concept for which they are intended (cf. Reeve, 1945). In reviewing over 200 suggested aids for the teaching of various concepts (ranging from basic numerical concepts to advanced concepts in algebra and geometry) we found only a few whose suggested use was based on more than just casual experience. Of the few which were evaluated by means of an objective index, adequate controls for comparison with other possible aids were lacking.

A number of investigators have pointed up the lack of clear evidence for the effectiveness of sensory-perceptual aids (Dodes, 1953; Lewis, 1956; Gibb, 1953; Fehr, 1947; Kinsella, 1950), and one (Dodes, 1953) suggested that there is no proof that any special lesson type, method, device, or classroom practice leads to improvement. It may be this lack of experimental work, which seems to be the exception rather than the rule, that leads to the heterogeneous results of a recent survey, in which attitudes were expressed ranging from enthusiastic support to extreme skepticism of the effectiveness of sensory aids (Syer, 1949).

We would, therefore, urge, as others have, continuing review, with carefully planned experiment, of the need and effectiveness of these devices, lest as Johnson and Nickerson put it, "...the development of future materials follow its expensive trial and error course." (1947).

REFERENCES

1. Allport, F. H., *Theories of Perception and the Concept of Structure*, New York: Wiley, 1955.
2. Barakat, M. K., A factorial study of mathematical abilities. *Brit. J. Psychol., Stastist. Sect.*, 1951, 4, 137-156.
3. Baravalle, H. von, Psychological points of view on the teaching of arithmetic. *Math. Teach.*, 1944, 37, 341-346.
4. Bartley, S.H., *Principles of Perception*, New York: Harper, 1958.
5. Berger, E. J., Devices for a mathematics laboratory. *Math. Teach.*, 1950, 43, *et seq*.
6. Blank, L., Teaching mathematics in technicolor. *Math. Teach.*, 1948, 41, 229-231.
7. Brown, H. Emmett, A science teacher looks at classroom films. *Sch. Sci.* and *Math.*, 1939, 39, 342-351.
8. Brueckner, L. J., Grossnickle, F. E., and Reckzeh, John, *Developing Mathematical Understandings*. Phila: Winston, 1957.
9. Bruner, J. S., and Goodman, C. C., Value and need as organizing factors in perception, *J. Abnorm. Soc. Psychol.*, 1947, 42, 33-44.
10. Carper, D., Seeing numbers as groups in primary-grade arithmetic. *Elem. Sch. J.*, 1942, 43, 166-170.
11. Cell, J. W., The trigonometric functions. *Math. Teach.*, 1950, 43, 187-192.
12. Commission on Post-war Plans., Second report of the commission on post-war plans - multi-sensory aids in mathematics. *Math. Teach.*, 1945, 38, 195-221.
13. Dantzig, Tobias, *Number, The Language of Science*. New York: Anchor Press, 1957.
14. Denbow, Carl, Teaching the laws of algebra, or mathematics is a language. *Math. Teach.*, 1941, 34, 360-363.
15. Dodes, Irving Allen, The science of teaching mathematics. *Math. Teach.*, 1953, 46, 157-166.
16. Fehr, H. F., The place of multisensory aids in the teacher training program. *Math. Teach.*, 1947, 40, 212-216.
17. Fehr, H. F., Theories of learning related to the field of mathematics. *Nat. Counc. Teach. Math.*, 21st Yrbk., 1953, 1-41.

18. Gibb, E. Glenodine, A review of a decade of experimental studies which compared methods of teaching arithmetic. *J. Educ. Res.*, 1953, 46, 603-608.

19. Gibson, J. J., *Perception of the Visual World*. Boston: Houghton-Mifflin, 1950.

20. Heidbreder, Edna, The attainment of concepts. VI. Exploratory experiments on conceptualization at perceptual levels. *J. Psychol.*, 1948, 26, 193-216.

21. Helmholtz, H. L., On the origin and significance of geometrical axioms. In Newman, J. R. *The World of Mathematics*, Vol. I, 647-668. New York: Simon and Schuster, 1956.

22. Johnson, D. A., and Nickerson, J. F., The future of visual aids in mathematics. *Math. Teach.*, 1947, Vol. 40, 180-181.

23. Judd, C. H., A psychological explanation of failures in high-school math., *Math. Teach.*, 1932, v. 25, 185-192.

24. Kinsella, J. J., Research in mathematics education. *Math. Teach.*, 1950, 43, 410-413.

25. Krech, D., and Crutchfield, R. S., *Elements of Psychology*. New York: Knopf, 1958.

26. Lewis, Eunice, The role of sensory materials in meaningful learning. *Math. Teach.*, 1956, 49, 274-277.

27. Lynch, J. M., Individual differences and course revision in plane geometry. *Math. Teach.*, 1942, v. 35, 122-126.

28. McCurdy, S. E., Colored chalk techniques for basic mathematics. *Math. Teach.*, 1948, v. 41, 369-371.

29. Meserve, B. E., Linkages as visual aids. *Math. Teach.*, 1946, 39, 372-379.

30. Miller, G. A., The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.

31. Newson, C. V., and Randolph, J. F., Trigonometry without angles. *Math. Teach.*, 1946, 39, 66-68.

32. Reeve, W. D., Multi-sensory aids in teaching mathematics. *Teach. Coll. Rec.*, 1945, 46, 420-427.

33. Smith, D. E., and Ginsburg, J., From numbers to numerals and from numerals to computation. In Newman, J. R., *The World of Mathematics*, Vol. I, pp. 442-464. New York: Simon and Schuster, 1956.

34. Syer, H. W., and Ingeneri, P. J., Multi-sensory aids in mathematics. *Sch. Sci.* and *Math.*, 1949, 49, 134-140.

35. Syer, H. W., and Johnson, D. A., Aids to teaching. *Math. Teach.*, 1948, 41 *et seq*.

36. Syer, H. W., Sensory learning applied to mathematics. *Nat. Counc. Teach. Math.*, 21st Yrbk., 1953, 99-155.

37. Wallach, H., and O'Connell, D. N., The kinetic depth effect. *J. Exp. Psychol.*, 1953, 45, 205-217.

38. Young, Edna H., Employing visual aids to teach locus in plane geometry. *Nat. Counc. Teach. Math.*, Eighteenth Yrbk., 1945, 109-116.

AN EARLY APPROXIMATION TO PRINCIPLES OF PROGRAMMING CONTINUOUS DISCOURSE, SELF-INSTRUCTIONAL MATERIALS*

Thomas F. Gilbert
Harvard University

I. INTRODUCTION

Any number of machines, gimmicks, and devices have been used in the attempt to cope with the problems of mass education. These mechanical contrivances have seldom played more than a minor, ancillary role in the classroom. The motion picture, phonograph, and radio have been the devices most frequently used, and television has recently developed considerable popularity (3). There are some who suggest that standard classroom procedures eventually will become ancillary to television modes of production.

Research on TV Education primarily has been addressed to the question, "can television teach?". The answer to this question requires no further research; the answer is clear. Television can teach and, generally speaking, teaches neither better nor worse than the usual classroom procedures (2, 7). We have been provided with no unequivocal answers to the questions concerning the superiority of TV or the classroom with one or another kind of pupil or subject matter. As Wischener (7) has pointed out, the failure to demonstrate any marked differences in the products of TV production and the usual classroom procedure reduces the question of TV implementation to administrative, logistic, and economic considerations. It is clear that the possibility of automation is, now, the most attractive feature of educational television.

The salient features of the educational problem are still before us. Automation of a set of operations that produce an inefficient product can only lead to more economical means for producing the same inefficient product. A covered wagon produced on the assembly line is still a covered wagon and will not bear comparison to a hand-carved Cadillac. What we need to automate are the events that occur when a good student learns. Whether television is the ideal instrument for this is an open question.

The relevant events that occur when an organism learns has been the primary concern of the behaviorists since the time of Pavlov and Thorndike. It is true that learning theory has been a matter of considerable controversy in behaviorism's 50-years history, and the controversies still rage. Yet, theorists have come to agree on a respectable number of principles established in the laboratory. We know a great deal more about learning in the laboratory than we have ever been able to engineer in the classroom; indeed, even the grossest principles of learning have received little attention in the reconstruction of the classroom. Major changes in educational methodology have been with reference to social and motivational properties, and have been particularly marked by the attempts, under the leadership of John Dewey, to make the schoolroom a situation somewhat more representative of the world outside the schoolroom. Reconstructions in the atmosphere of the classroom may have relevance to the generalization of what is learned outside the classroom, but they can have very little relevance to the actual acts of original learning. If no learning occurs the question of its generalization is a purely academic matter. Of course, John Dewey was concerned with far more than the atmosphere of the classroom. It was the central theme of his philosophy, from his ethics to his logic, that behavior is shaped by its consequences. It required the development of systematic learning theory, particularly by B. F. Skinner (6), to capture the richness of Dewey's functionalistic program in the detail of moment to moment changes in behavior. In the absence of the observations of the laboratory, the translations of his program into detail was left to who-would, and some of these translations have been indeed grotesque, as Dewey lived to observe. Laboratory efforts have led us to induce some principles of behavior with a precision not possible with the methods of philosophy. Whereas these principles typically developed out of the animal laboratory, procedures for their engineering deserve a laboratory of their own. We are concerned here with just what these principles are and what procedures are optimal for engineering learning.

---

* Part of this report was prepared while the writer was employed at the University of Georgia, and part of the report was prepared for the Bell Telephone Laboratories.

A. Behavior Principles and the Classroom

As in any science, the principles of behavior are stated in a technical language that necessarily loses precision and force when translated into ordinary terms. Nevertheless, these principles are stated here in non-technical form to provide a basis for understanding why the semi-automation procedures have developed as they are.

1. The organism is conditioned not by mere repetition of the acts (responses) to be learned, but by certain consequences that follow from these acts. The consequences that are found to strengthen these responses are called reinforcers (less technically, rewards). This principle suggests that very little learning can occur through watching someone else perform the task to be learned. The educational system is not devised on the assumption that mathematics and psychology are spectator sports. In learning tennis we may be oriented to the task by watching the tennis instructor hit a few balls. Similar orientation can occur in reading a manual on the subject. The skill is developed by the student, however, in the process of his actually working at the task. In short, behavioristic principles underline the age-old truism that we learn by doing.

It is clear that the student must perform the acts to be learned, and that suitable consequences must occur to select out and strengthen these acts.

2. Reinforcers strengthen responses that appear immediately prior to reinforcement. The greater the delay between the occurrence of the response and the reinforcer, the less likely is learning to occur.

The immediacy of reinforcement is, perhaps, the most important single variable in learning. The principle is frequently violated. Left to his devices, the student will attempt to perform some of the acts involved in the task to be learned. Seldom is there an opportunity for him to receive the immediate knowledge that he is correct or not. As a result of the lack of expertly devised immediate reinforcers for correct performance, we can only guess that the student supplies his own brand of reinforcement. He may tell himself that he is probably right even after incorrect performance. The fact of making progress in completing a task, even if part of the performance is wrong, can be reinforcing. In this case, the student is being reinforced for incorrect performance. All major learning theories would agree that task completion can be reinforcing, and one theory proposes that *all* reinforcement is of this nature. At any rate, the delayed reinforcement occasionally supplied by the teacher (24 or 48 hours after the task) must compete with the immediacy of reinforcement supplied by task completion.

3. *Repetition of a response without reinforcement weakens a response.* In the absence of skillfully devised immediate reinforcement, whatever tendencies the student may have to perform correctly may drop out of his repertory.

4. *"Right" responses can be differentiated from "wrong" responses by reinforcement of the one and non-reinforcement (more technically, extinction) of the other.* (Punishment is a complex matter, and is, in no true sense, the opposite of reinforcement.) As was described above, the student, in supplying his own reinforcement, will, inadvertently, strengthen incorrect performance. Teacher supplied reinforcement may also, unintentionally, be of a non-differential nature. The "80" a student receives may seem too good for him to bother with the details of what, specifically, he did correctly or incorrectly.

5. *Motivation is produced by reinforcement. Within limits, the more frequently an organism receives reinforcement, the greater the motivation to work. The intensity of any given reinforcement has little or no effect on the rate of learning.*

A great part of teacher time is spent in attempts to motivate students. This, in itself, violates no behavioristic principles, although it certainly creates a number of problems. Although it is questionable that a high degree of motivation will, itself, facilitate learning, it certainly facilitates the performance of responses already in the organism's repertory. If study habits are bad, it will be the performance of these habits which will be facilitated. Every teacher has observed the energetic student who seems to acquire error more rapidly than truth. A more important consideration is what kind of events contribute to student motivation. The findings of the behavior laboratory point to an inescapable conclusion: reinforcement is the primary source of motivation. This is technical confirmation of the old saw, "Nothing succeeds like success."

A typical result of too little planned reinforcement is the student who seems to have no "get up and go." In the laboratory, infrequent reinforcement for a considerable amount of work can produce such extremes of motivation loss that the hungry animal may crouch in the corner seemingly taking no interest in anything. The homologue outside the laboratory is found in the student who reports that he is "discouraged" or "just can't get going." This state is produced strikingly in the student who has a long task before him with little assurance of success. If many tasks in a situation require long periods of work in order to obtain reinforcement, the motivational loss may generalize to the entire situation. The student reports losing interest in "everything." The high frequency of this attitude, as well as observations of classroom procedures, lend strong support to the conclusion that many of our educational ills result from inadequately planned reinforcement.

6. *The conditioning of a complex skill is done most effectively by "shaping." In the beginning, any responses in the organism's repertory that resemble the skill are reinforced, on successive occasions reinforcement is given only to those responses that more closely approximate the skill.* There

can be no argument about the frequency with which this principle is violated in the schoolroom. Students typically, if not nearly always, attempt to master higher level skills before they have mastered earlier and more fundamental material. Pleased with his "80" as a signal of satisfactory performance, the student proceeds to more difficult material with the understandable encouragement of the teacher who may find the "80" good by comparison.

These six propositions certainly do not deplete the list of important behavior principles. They do seem somewhat more remarkably relevant to the initial needs of educational engineering. The teaching programs described in this report are no panacea for educational ills. These principles can be, and have been, put to use by teachers since time immemorial. In some important ways, the programs can carry out these principles more efficiently and more inexpensively than is commonly the case, freeing the teacher from much simple routine. Rather than replacing the teacher, the programs should provide him with time to utilize his talents in more important ways.

B. Recent Attempts to Apply Behavioral Principles to Teaching

We have indicated that one of the important principles of learning is the utilization of immediate reinforcement or the immediate feedback to the student that will allow him to evaluate his performance at any given stage of learning. This immediate confirmation could undoubtedly be arranged by his teacher in the classroom had she but one or two students before her. When she presented a problem to the student who made some response, she could tell him immediately whether he was right or wrong. The disadvantages of this procedure are economical in nature. In the first place, it is unreasonable to have a teacher working with one student at a time, since this procedure would require as many teachers as there are students. Secondly, it would be a demeaning task for the teacher, particularly with simple subject matter, to stand making such confirmatory remarks to the student. Various mechanical devices have been created which will perform this task in lieu of a teacher. These devices have gone under the generic name of "teaching machines," though the name is somewhat unfortunate since the machines themselves are constructed with one purpose in mind, and that is to provide immediate confirmation to a student following his response to a problem. "Automatic teaching device" is a preferable name, since some of these devices are not machines at all. Automatic teaching devices are no magical contrivances, and can be no better than the material that is put into them. It is the analysis and sequencing of the problems that should form the core of research on teaching programs.

Automatic Devices for Continuous Discourse Teaching. The history of teaching machines and their generally unsatisfactory performance recently has been described by Porter (3). The psychologist Pressey (4) devised what probably has been the best known of these machines. It consisted of a device resembling a child's toy cash register. A problem appeared in the window of the device along with five multiple choice answers, one of which was correct. The student selected what he considered to be the correct answer and pressed the corresponding key. The machine would then let him know whether he was right or wrong. Although this device was never adequately tested, it did not appear to be a glorious success (3). For one, the requirement that the student recognize a correct answer from a small group of possible answers violated at least two principles of modern learning science: (a) The student was learning to recognize one of the restricted group of alternate responses rather than to produce the correct response himself. The homology in sports would be the grandmother who is an ardent baseball fan. She knows a well executed hit and run play when she sees one, although she has never wielded a bat. (b) The student has little opportunity to produce erroneous responses already in his repertory, and these, in turn, have little opportunity to undergo extinction. They remain to compete with whatever recognition habits he may learn by the machine.

Recently, Skinner has turned his attention to applying behavior principles to classroom teaching in the belief that the art of teaching has not kept pace with the science of learning (6). Before describing the specific procedures he has developed, I shall describe a somewhat idealized set of procedures, illustrating how the principles are applied. The several specific procedures are practical compromises with the theoretical ideal.

This machine, which exists in only crude design, is a box about two feet square and nine inches high. A disk with a number of problems is placed in the machine. One problem is exposed under a glass window. In an open window there is exposed an inch or so of adding machine tape. The student writes on this tape what he believes to be the solution of the problem. He then pushes a button that advances his answer to a position underneath a glass window. He cannot retrieve his answer, nor can he erase it. The button that performs this operation also releases a lock on a device resembling an Address-O-Graph. This device is in the upper left hand corner of the machine. A second button is pushed that opens the checking device to the appropriate place, exposing a set of 12 answers, one of which is correct. Each answer has a number. The student selects the number of the solution that matches his (if any do), and sets a selector dial to that number. At this time the checking device closes, and if the student's solution is correct a buzzer sounds and a new problem is presented in the problem window. Also, if the answer is correct, a hole is punched in the adding machine tape, recording the fact of a correct solution. If the student's solution is wrong, the tape remains in position, the problem remains in its window, and the checking device remains locked. He cannot proceed to a new problem until he has provided the right solution, so that he begins again and writes in a new answer.

This machine allows the student to work at his own rate, and prevents him from randomly setting the selector dial to various positions until he hits upon the correct combination. These two features are provided for in the following way. Suppose that the machine has in it a number of problems in factoring algebraic expressions. A switching circuit in the machine is so designed that factoring problems keep appearing in the machine until the student has given five consecutive solutions on the first attempt. The chance of obtaining five consecutive correct selector switch settings on the first trial is, in this case, about one in 25,000. A student who tried to beat the machine by making random selector switch settings would have factor problems repeated to him during the allotted time. When the student actually comes to understand the principles of factoring well enough to give five consecutive first trial answers, the machine automatically drops out the factoring problems and begins to present the next higher level material. The machine literally "senses" the fact that the student has induced the principle of factoring.

Let us examine how the machine is applying the principles set forth above. First of all, it is providing reinforcement for every correct response. It is providing this reinforcement immediately. It does not allow reinforcement of wrong responses. It requires the student to master habits at one level before proceeding to the next. Most important, the student does not watch someone else perform the tasks—he must do them himself. Furthermore, the machine is not restricted to "rote memory" materials. In the example given above, the student must induce a principle or new factoring problems are constantly presented.

One of the advantages given the machine is that it has only one student to teach at a time. The machine does nothing a teacher could not do if he had only one student (no electronic circuit can be devised that is as nimble as a human). As Skinner has pointed out, if the teacher had only one student it would be a waste of his talents to provide reinforcement such as "yes, 2 + 2 is 4". The machine can do this more inexpensively than a human teacher. And, surely, the teacher has better things to do.

One principle, that of maintaining motivation by high frequencies of reinforcement, would not be entirely satisfied by the machine as it has been described. If a given problem in the machine is so difficult that it requires large amounts of work with no reasonable assurance of early success, the student would suffer from the same motivational loss as described above. Skinner has attempted to meet the requirements of the high frequency reinforcement principle by presenting a method for designing the material. Each class of materials is analyzed into the component parts of its development. Beginning with the elements of the material most likely to be already in the student's repertory, the problems gradually become more complex and come to be stated in the appropriate technical language step by step. This assures a greater frequency of reinforced responding. This procedure illustrates the principle of shaping, described above.

Skinner's Existing Machines. Skinner has devised three general machines. The first is a pre-school machine that is designed to aid children in the readiness program. It is at such an early stage of experimentation that it would be useless to describe it here. The second machine has most of the features of the idealized machine described above. It is constructed primarily for the purpose of teaching children arithmetic. This device, resembling an adding machine, has a large window into which problems can be presented. A set of sliders allows the student to arrange different numbers, letters, and arithmetical operators into an answer window by simply placing the appropriate sliders in position. A crank on the machine, when turned, will present a new problem only if the student has provided a right answer. This machine never shows the student an answer, but only tells him whether he is right or wrong. This device presently is being manufactured by IBM and may soon be commercially available.

A third machine has been constructed for use with students at the high school and college level. This consists of a box about the size of a portable phonograph. Problems are placed in the machine on a disk about the size of a phonograph record. The machine works much like the machine described above except for one important difference. Every time the student writes his answer on the tape and pushes it under the glass, the correct answer is exposed. If the student's answer is correct by comparison with the machine-provided answer, he pushes a button that not only marks the answer tape but also drops the problems out of the machine. If he is in error, the problem reappears on the next rotation. The dropping out of a problem is achieved by means of a detent underneath the turntable. Each problem's position has a detent that can be moved out of position by the button that marks a correct answer. The adult level machine is completely mechanical and has no electrical connections.

Other Machines. Douglas Porter, in the School of Education at Harvard, has devised a machine that is ingenious for its simplicity. Having only one moving part, a crank attached to a roller, the machine will take a sheet of ordinary size writing paper. Problems written on the paper can be cranked through one at a time. Once the paper is started it cannot be reversed. The problems appear in a plexiglass window, and the correct answer can be seen only after the student's answer is pushed forward under glass. The answers are written on the sheet on which the problems are dittoed. The major disadvantages of Porter's device (which, by the way, costs only $20) is that so few problems can be put on any one sheet. Someone must supervise the student to keep him from seeing the problems and their answers on the sheets that have not been put into the machine. Furthermore, the problems will not drop out once they are answered correctly.

The writer has designed a machine in the attempt to combine the features of Skinner's college machine with the simplicity of Porter's device. It consists of a box 21 inches long, 16 inches wide, with a surface sloping toward the student from a ten inch height in the rear to two inches in the front. The machine has a problem window and an answer window. The problem window is covered with glass, except the window in which the student writes his answer. The problem window is partially covered so that the correct answer will not appear until the crank covers the written answer with plexiglass. A shutter, operating with the crank, will uncover the correct answer so that it may be matched with the student's answer. Inside the machine is a roll of adding-machine paper on which the answer can be written. Materials with problems and answers can be inserted into the machine. The machine is locked so that the student cannot get inside to change an answer. A button will mark a correct answer. This device combines much of the simplicity of Porter's machine, with all but one of the features of Skinner's machine. It will not drop out problems.

Several people, such as Ferster (1), have made an even simpler masking device which can be used with ease. One such device, constructed by the writer, consists of two file folders pasted together, end on end, and a problem and answer window cut into one of the folders. The answer window is covered by a flap. Mimeographed materials are pushed through the folders by the student. He writes his answer on a sheet of notebook paper and lifts the answer-window flap to check his answer. This device has been tested in the classroom and has proved, so far, to be quite successful. The folders do not last long, the student can "cheat" the device, and it cannot be used for serious testing purposes. Otherwise, it can be quite helpful, requiring rather little supervision.

II. THE CONSTRUCTION OF TEACHING PROGRAMS

The following rules for the writing of teaching materials represent only a crude approximation to optimal principles of programming. These principles arise partly from the framework of behavior analysis discussed earlier, partly from the writer's experience in writing teaching programs, partly from observing neophytes learning to write programs, and partly from the experiences of others who have written programs. About all that can be said for their validity is that programs that were written according to them have been demonstrated to teach with considerable success and, further, have a certain degree of plausibility. The reader must be cautioned that, at this stage, the explication of these principles represents a highly personal document. It is an attempt by the writer to describe his behavior while writing programs. The behavior, even if accurately described, is not likely to be the last word in wisdom. It must be remembered that it is the essential problem of a research program in teaching to make a more exacting approximation to principles of programming. The following rules will, then, serve as a baseline from which such research can proceed, and as an instructional guide to those who are writing programs under the present scheme.

A. Designing Items to Evoke Relevant Behavior

The first principle pertaining to the construction of materials is that the subject matter should be analyzed into its stimulus and response components. A person who constructs material for the machine has to be aware of what responses to what stimuli are made by a student who efficiently develops skill in the subject matter. This requires analysis of the behavior of, say, an elementary algebraist, or an elementary psychologist. There are two problems of relevancy here: a) The problem (or stimulus) that appears in the window of the automatic teaching device should represent the kind of situation to which you want to teach students to make responses. b) The responses required of the students should represent the kind of response you want him to make. If the response you're requiring of him is already in his repertory, or is not the kind of response that the task to be learned typically calls for, then such a requirement is irrelevant. As an example, suppose we were teaching business office service-representatives in the telephone company to read and interpret telephone toll tickets. Imagine the student seated before a teaching device holding appropriate materials and having a set of toll tickets, filled in, next to the device. The material listed below would represent, in the left hand column, the kinds of problems presented in the problem window of the device; and the materials in the right hand column are the answers to the problems.

ITEM	ANSWER
1. Exhibit one is a toll ticket. It is a record of all long distance calls. The information on it is filled in by the operator at the time of the __________.	call

2. Exhibit one is a (a)______. The date of (b) ______ call is placed at the top of the ticket.	(a) toll ticket (b) long distance
3. Exhibit one is a ________________.	toll ticket
4. The date of the (a) ______ call is placed at the (b) ________ of the ticket.	(a) long distance (b) top

The responses required in item 4 are irrelevant. Items to follow will require the student to discover the date at the top of the toll ticket. Teaching him repetitively to say "the date is at the top of the ticket" is unnecessary since he will practice discovering it there. Furthermore his ability to see where the date occurs is not a requirement of the task. He will later be able to say "yes" by virtue of his practice in discovering the date.

The response required in item 1 is also irrelevant to the task to be learned. It is, however, an exception to the rule of irrelevancy in that the response cannot be easily made unless S has read the item. The requirement of the response helps insure that S reads the informative statement.

Response to 2a is not irrelevant since it is desired that S be able to designate the ticket as a toll ticket. Further practice will be necessary since S must recognize a toll ticket in order to perform later problems. Item 3 is, therefore, unnecessary. 2b is justified on the same grounds as 2a; it requires S to note that toll tickets on long distance calls are related. Further practice is again unnecessary since further practice with toll tickets will have to do only with long distance calls. Item 2 prepares the student for actually reading tickets, and this practice should begin immediately. We want to teach S to read tickets, not say things about them. Therefore, it is important that items requiring actually reading the tickets should begin as soon as possible. Regard a few more examples from the Service Representative Training Program:

ITEM	ANSWER
5. See ticket No. 4. What was the date of the call? (a) ______. What was the date of the call on ticket No. 4? (b) ________.	(a) Feb. 16, 1958 (b) April 3, 1958
6. No. 9 designates the date on which (a) _____. This date is (b) ______.	(a) a long distance call is made. (b) Dec. 1, 1954
9. The date of a long distance call is at the _______ of a toll ticket.	top

Items 5 and 6 call for two different kinds of responses. 5 requires S to supply the date to a question that calls for it. 6 requires S to interpret the date upon seeing it. The stimulus in item 5 becomes a response in item 6 and vice-versa. This is justified if both responses are relevant to the task. Item 9 is unnecessary since it requires S to make an unnecessary statement describing an act (looking at the top of the ticket) he has already learned to perform. Items 5 and 6 represent an example of stimulus-response reversal. Such a reversal should follow the principle of relevance, the interchange taking place only when both responses are called for in the task. Another example of stimulus-response reversal can be supplied from a program teaching the color code for resistors. In this segment of the program you may assume that the code has already been learned by S and he is practicing its use with a group of resistors mounted on a panel by the machine.

ITEM	ANSWER
21. Look at resistor No. 16. What is its value?	72,000 ohms
22. What resistor has the value of 29,000 ohms?	No. 31

Another example of stimulus-response reversal can be provided from a program in algebra:

ITEM	ANSWER
41. The commutative law for addition states that ____________.	a + b = b + a
42. a + b = b + a is an example of the _________________.	commutative law for addition.

Stimulus-response reversal should, however, follow the rule of relevance, and should not be made uncritically.

In the experience of the writer the most common tendency to create irrelevant items or items calling for irrelevant responses can be illustrated by item 9 above. Item 9 is unnecessary since it requires S to make an unnecessary statement describing an act (looking at the top of the ticket) he has already learned to perform. It often is easier to write items requiring students to make verbal descriptions of acts they are to learn rather than to write items that require the students to make the acts themselves. In item 9 you do not want to teach the student to say that the date is at the top of the toll ticket, but you want to teach him to look at the top of the toll ticket and find the date. The item should, then, require him to do just that. Item 9 would better read:

ITEM	ANSWER
9. The date of the long distance call on ticket No. 16 is ___________.	May 19, 1957

There are two exceptions to the rule of relevancy. One of these exceptions would allow the item writer to require, under certain circumstances, a verbal description of an act to be performed in lieu of that act. Suppose that you want to teach a telephone business office representative that she is to answer the telephone always on the first ring. Using only written materials and a Skinner type teaching machine, it would be difficult, in the absence of an instructress, to provide confirmation of a correct response of this nature. An item may read something as follows:

ITEM	ANSWER
2. To answer the telephone promptly is to answer on the ______ ring.	First

Here you want to teach the representative that promptness means answering on the first ring. Instead, she is taught to say that promptness means to answer on the first ring. However, this verbal behavior is so uniquely related to a skill already learned by the representative that to require more than this in the program may be unnecessary. After all, one of the great values of verbal behavior is that it is a mechanism for conveying the course of completion of non-verbal behavior. When this mechanism is well established in the student's repertory, then the verbal behavior may be essentially interchangeable with the acts that they indicate. Danger in relying on verbal responses occurs on those occasions when the verbal response is not related to some other well developed skill. When the verbal response is interchangeable with a skill that is being taught, then the requirement of a verbal response is not, after all, violating the principle of relevancy.

A second violation of the principle of relevancy may be valuable in those instances where it is desired to present an item which is designed to give information and which does not require a response. Informative frames are, however, likely to be read by the student too quickly and too inattentively for him to grasp the information. By merely deleting one or two words from the informative frame, requiring him to fill them in, you may force him to pay very careful attention to what he reads, since he must read carefully in order to fill in the missing words. In these instances it does not matter what the missing words are except that anyone who has read the material should be able to supply them, and they should be difficult to fill in unless the student has read the material very carefully. An example calling for this type of "irrelevant" response may be supplied from the algebra program. We want to inform the student how to multiply a monomial times a binomial:

ITEM	ANSWER
56. To multiply a monomial times a binomial one ______ each term of the binomial by the monomial.	Multiplies

In this case stimuli for the required response are clearly in the item and anyone who reads the item should be able to supply the word. Incomprehensible reading of the item, however, will result in failure to make the response. This technique, suggested to force careful reading on the part of the students, is a questionalbe one. Our experience has shown that students generally discover that inadequate absorption of information in one frame will lead to failures in making correct responses in succeeding frames. This seems to be sufficiently aversive to them to lead to more careful reading. Another instance in which requiring nonessential responses may be of value is where information is usually provided when it is of too great a quantity to insert in the frame itself. When the items to follow are dependent upon the students having read the panels carefully, a few items may be inserted at the beginning which will question him about the panel itself. These items will be of such a nature that he could not correctly respond to them unless he had carefully read the panel; thus the items force him to read the panel before going on to the actual teaching items.

The program writer has the further problem of arranging for contextual relevance. When teaching a student to respond to such problems as $a + b = b + a$ as the commutative law for addition, the problem should also be stated in other symbols, such as $k + m = m + k$. In teaching trigonometry, a student is required to make certain descriptions of the characteristics of triangles. These triangles should be presented in various positions lest generalization fail to occur. In teaching the color code for resistors, supposing the resistors to be mounted on a breadboard next to the teaching device, the student should have to respond to resistors which are arranged in various positions, some projecting out from the board, some having to be read left to right, some right to left and some up to down. Part of the analysis of subject matter will require the writer to state the class of relevant problems, rather than single instances, and provide a sufficient number of single instances so that responses to the class of stimuli can be made. In repeating items it is therefore well to repeat the items in different forms. For example, in the algebra program we want to teach a student to recognize the fact that numbers are represented by letters. Various items directed to this and spread throughout the algebra program would read somewhat as follows:

ITEM	ANSWER
1. The algebra letters stand for______.	Numbers
2. The letter X is a symbol used for a ________________.	Number
3. The letter B is a symbol used for a ________________.	Number
4. If letters like X, Y, A, B etc. represent numbers they should be added like ___________.	Numbers
16. In the statement of this law, the letters A and B represent __________.	Numbers
17. Letters like R and J do not stand for pigs and horses, they stand for ____________.	Numbers
57. The statement $(f + c) + d = f + (c + d)$ is called the associative law for addition. Your experience tells you it will be true no matter what ____________ f, c and d stand for.	Numbers

---

In these problems we have one purpose in mind and that is to teach the student to recognize that letters, in algebra, stand for numbers. To this end we have provided a number of contexts in which a problem is presented.

An additional problem in arranging for relevancy is the necessity for having a system whereby all responses equivalent to correct answers in the program can be confirmed. The following items selected from various programs will exemplify this:

ITEM	ANSWERS
1. Customers are given ________ attention, are not kept waiting.	prompt, immediate, (or equivalent)
2. The representative uses a contact memorandum to _______ during each customer contact.	makes notes (or equivalent)
3. $d(ab + c) =$ _________.	dab + dc, or abd + cd, (or equivalent)
4. The muscles chiefly involved in chewing are ___________.	Striped (Skeletal)

Care must be taken in examining the programs to determine whether the students are too free in judging their answers to be equivalent to the correct answer provided.

B. Problem Repetition

It is desirable that any particular problems, or kinds of problems, should be repeated until the student's correct responses to them have achieved a high degree of probability of occurrence, and a minimum fading of this probability with time. Five conditions are desirable:

a) A minimum number of problems of a class.
b) A minimum sample of kinds of problems in a class.
c) A minimum time invested per student.
d) A maximum probability of correctness of response.
e) A minimum fading of probability during a stated time.

Particulars of a given class of problems, as particulars of any other class, will vary along several stimulus dimensions. The minimum number of necessary, particular problems will be at least two and probably a great number more. For example, the following factoring problems differ in several respects:

1) $A^2 + AB^2 = A(A + B^2)$

2) $7XY + 14X^3Y - 28XY^4 = 7XY\ (1 + 2X^2 - 4Y^3)$

These problems differ in the size of exponents, the number of terms, characteristics, notations, use of numbers as terms, and probably other characteristics. It is desired that the student generalize the operation of finding prime factors for any set of notations, then, at least two problems differing in notation should be presented. Repetition of problems should be different in as many particulars as possible.

It seems plausible that repetition of problems should form a kind of density gradient, such that in the introduction of a new concept a number of problems are given close together and that, later, repeated practices of these problems be spread more thinly throughout the program. This is just the old principle of distributing practice.

Certainly no exacting statements can be made about the numbers of repetitions required for certain kinds of problems. However, certain general guides can be suggested. For a particular type of material, some items will serve as building blocks for the accomplishment of more complicated material. For successful responses to the more complicated problems, less repetition of the more simple problems is necessary, since the simple problems will have to be solved over and over again in the solution of the more complicated ones. For an example, in teaching the multiplication and division of sign numbers, after sufficient problems have been given to induce the rules of multiplying and dividing sign numbers, no further specific repetition is necessary. In problems that will follow in the algebra course, utilization of the rules of multiplying and dividing sign numbers will be required. Where problems stand more or less to themselves, such as translating into numbers the color bands on the resistor, or spelling words, more specific repeated practice trials probably should be added. Where responses are rather arbitrarily linked to their stimuli, such as again translating into numbers the color bands on a resistor, a great deal more repetition will be necessary to insure delayed recall. The amount of repetition of arbitrarily linked paired associates can be cut down by introducing mediating associations, or prompts, that will make the connection seem less arbitrary. For an example, in teaching the color code for resistors, we provided a mediating association for each color-number pair. The items appeared something as follows:

ITEM	ANSWER
1. A black band on a resistor stands for 0: 0, black, nothingness. The black band stands for ________.	0
2. A brown band on a resistor stands for the number 1: 1, brown, penny. A black band stands for ________.	0
3. A number 1 is represented by a (a) ______ band, a number 0 is represented by a (b) __________ band.	(a) brown (b) black

---

Since mere repetition of problems can be rather aversive in learning, it is often well to combine the repetition with new information in a given frame. The term interlocking will be applied to this procedure. Suppose that we have taught the student that static electricity is produced by friction between objects causing electrons to be dislodged from their orbits. We may wish to give him further practice in saying that friction produces a static electricity, and then to move on to the notion that when an atom loses an electron from its orbit, it becomes negatively charged. The following item would represent a way in which the repetition can be combined with new information.

ITEM	ANSWER
1. Static electricity is produced by the _______ between objects. If the atoms in a material have the electrons dislodged from their orbits, the material is then said to be negatively charged.	Friction

---

Another example from the algebra program is as follows:

ITEM	ANSWER
1. The statement (A + B) + C = A + (B + C) is called the associative law for addition. Your experience tells you that it will be true no matter what ______ A, B, C stand for.	Numbers

Here we provided for a repetition of making the response that numbers are represented by letters in algebra, as well as presenting the statement of the associative law for addition. Notice that in both of the above items the repetition of already learned material and the introduction of new information are done in such a way that they do not interfere with each other and stand together as a meaningful unit.

Another principle of repetition might be called dovetailing. This is used when one must learn to make a set of different responses to different, but related, stimuli. For example in teaching algebra, we first introduce the commutative law for addition, following which the associative law for addition is introduced. After some repetition with the commutative law, the associative law is introduced with some repetition, and then the two laws are presented together, so that further repetition is done in such a way that the appropriate responses to the two sets of stimuli will not be confused. In a general way, this would indicate teaching the first stanza of a poem to some degree of mastery, following this the second stanza of a poem to a degree of mastery, and following this practice with the two stanzas together. In teaching the color code for resistors, first teach that black stands for zero, then that brown stands for one, and then while requiring responses to both of these in the same frame, we introduce a new color-number combination, red: two.

In summary we have found that the aversive characteristics of repetition can be reduced considerably by varying the context of each additional problem, by introducing new information related to a particular problem, and by requiring discriminations to be made between two or more problems separately accomplished in the past. In any case, the aversive properties of repetition have not been overwhelming. It may be necessary in the future to experiment with programs having varying amounts of repetition in order to discover some more exacting rules concerning this problem. It has also been suggested many times that it may be wise to write different programs for students at different levels of capacity. Thus bright students would not be burdened with excessive amounts of repetition. Answers to these questions, as well as more precise statements as to what the questions are, can be discovered only in the teaching laboratory.

C. Augmenting and Fading

Most subject matter has continuity to it and should be taught as such. Programmed teaching differs decidedly from test instruction although the failure to make this distinction is very common among novices. A first attempt at writing a program produces items that look like test items. An algebra program written by a novice may begin something like this:

ITEM	ANSWER
1. The associative law for addition states that ___________.	A + (B + C)
2. The communitive law states that ___________.	A + B = B + A

A program beginning like this is not teaching since it is requiring responses of a student that must be obtained elsewhere. A teaching program, like any teacher, provides information that requires responses of students to insure their understanding of the information, and then requires utilization of the information by the student. It guides him from one step to the next, assuring that he has mastered one step before going to the next. It is the informative nature of teaching program items that marks them off most radically from testing items. A second characteristic of teaching items that differentiates them from testing items is their continuity, in that one item leads into another. While there are some sequential characteristics to many tests (that is they tend to go from the easy to the difficult material), their major characteristic is that each item is independent of the other. Having answered one item does not make it easier for the student to answer another. The opposite is true with teaching program items.

In the writing of a teaching program, you begin by making rather stringent assumptions about the information possessed by the dullest and most poorly informed student you wish to teach. When teaching electricity, you do not assume that he knows elementary electricity simply because he has had a course in the subject. Having decided upon this minimum information, you begin writing your program at that point. There are two general ways in which the guidance through the program is created. One, is by augmenting, or introducing bits of information that will lead the student into a new concept. An example of augmenting may be shown in a psychology program:

ITEM	ANSWER
47. In training a dog to salivate to a bell, we sound the bell just before putting lemon juice on the dog's tongue. The lemon juice elicits the saliva, but after repeated pairing of the bell with the lemon juice, the bell eventually comes to ______________ the salivation.	Elicit

ITEM	ANSWER
48. If we say that the bell is a condition stimulus (CS) and the lemon juice is an uncondition stimulus (US), it is clear that the CS is being substituted for the ________ to form a new reflex.	US (lemon juice)
49. In the conditioning of a response to a CS, the CS is being (a) ________ for the (b) ________ to form a new reflex.	(a) substituted (b) US
50. This kind of conditioning is said to be stimulus substitution or type S conditioning. The S in type S obviously stands for ________.	Stimulus substitution
51. In type S conditioning, the (a) ________ is substituted for the (b) ________ to form a new reflex.	(a) CS (b) US
52. What happens in type S conditioning? ________________.	One stimulus (CS) is substituted for another (US) to form a new reflex.

The above represents some items (moving much too rapidly, but for the sake of example) that bring the student from the situation of pairing a bell with lemon juice to a technical characterization of type S conditioning. During the course of this he was being provided with new information as well as being required to repeat and put to use the old information. Another example from the algebra program is given with some of the steps deleted for the sake of a shorter example:

ITEM	ANSWER
5. 3 + 2 = 5, also 2 + 3 = ________.	5
6. But also 7 + 10 = 17 and 10 + 7 = ______.	17
7. We can say that 3 + 2 is equal to 2 + 3 and 7 + 10 is equal to 10 + ______.	7
8. 7 + 10 is equal to _____ + _____.	10 + 7

ITEM	ANSWER
11. We can say that this will always be true by saying that a + b is equal to b + ______.	a
12. The statement that a + b is equal to b + _____ is called the commutative law for addition.	a

Here, by providing the student with examples of addition reversal, we have led him gradually into the statement of the commutative law for addition. By following such a scheme, he can go from the beginning of the algebra program, in which the student is taught that letters stand for numbers, to the end of the program, in which he is performing problems in quadratic equations, without his ever having to make a mistake, provided that the augmenting is proper and the items are unambiguous.

Fading or vanishing refers to the gradual withdrawing of information so that the student must more and more rely on himself. An example from the algebra program is as follows:

ITEM	ANSWER
14. The commutative law for addition states that a + b = b + ______.	a
15. The commutative law for addition states that a + b is equal to ___ + ___.	b + a
16. The commutative law for addition states that ____________.	a + b is equal to b + a

---

Another example from a sixth grade spelling program provided by the following items:

ITEM	ANSWER
41. Circle the word that may mean to be in pain: suffer, swell, improve, drug, fountain.	suffer
44. Write in the missing letters. They are both the same: I hope Queenie does not have to su__er.	suffer
47. Write the missing letters: I hope Queenie does not have to s_f__r.	suffer

ITEM	ANSWER
49. Write the missing letters. I hope Queenie does not have to s____r.	suffer
51. Write the missing letters. I hope Queenie does not have to s_____.	suffer

Another example from the psychology program is given by the following item:

ITEM	ANSWER
61. In type S conditioning, the CS is substituted for the US to form a new ________.	reflex
62. In type S conditioning the ________ is substituted for the US to form a new ____________.	CS, reflex
63. In type S conditioning the _______ is substituted for the _________ to form a new ________.	CS, US, reflex.
64. In type S conditioning _________.	the CS is substituted for the US to form a new reflex.

In each of the above examples, the prompts given by the teaching program are gradually withdrawn and the student is required to supply his own prompts.

Just how gradual should the augmenting and fading be? Augmenting can be quite gradual or quite rapid. It can consist of the repetitive introduction of small bits of information before new bits are presented to the student. Or it can be characterized by the presentation of very large amounts of information in successive frames. Similarly, fading may vary in graduation by having all the information that occurs in one frame removed for the following frame to the other extreme of continued presentation of the information with very small amounts of it withdrawn on successive presentation. Naturally these two characteristics of programming, augmenting and fading, are not entirely independent of each other. The repetitive presentation of small amounts of information is roughly equivalent to the gradual withdrawal of small amounts of information from a frame involving a great deal of information. The rules that we have followed in writing teaching programs to date are based on very limited experience with variations in graduations of guidance through the material. The primary rule that we have been following is this: The program should be advanced by such small steps that it is virtually impossible for a student to make an incorrect response. In following this rule, one would then write program materials in such a way that any errors in graduation would be on the side of making the steps too small rather than too large. The assumption implied by this rule is that the optimal ratio between errors and correct responses is zero. In the absence of any other information to guide us, this rule grown out of the motivational principle that motivational loss is produced by lack of reinforcement. It is clear from our experience with teaching programs that if too many errors are made by the student, the program becomes aversive (such as the teaching of the classroom) and the student tends to quit work. We do not know what the optimal ratio of errors to correct responses may be and, furthermore, we do not know whether this ratio changes from the early stages of a program to the later stages. That is to say it may well be that the optimal program will allow few errors in the beginning and an increasing number of errors toward the end of the program by making the program successively more difficult. Answers to these questions could only be discovered in the laboratory. This rule, as the other rules, is a tentative one dictated by pedestrian experience and whatever "wisdom" that has been provided by experiences in the behavioral laboratory. We have observed that, even in the best program, a certain number of errors will be made by the student. These errors will be produced by inattention on his part in reading the frames. The present belief is that such errors are quite beneficial to the course of the teaching program. This is particularly so since the student by these errors learns to pay careful attention to the frames.

Whereas the primary reason for the gradual guidance of the student through the materials is the development of the subject matter itself, a secondary reason is the maintenance of a student's motivation to work at the program. Very fine guidance has appeared to maintain a rather high level of motivation in students, and this fine guidance like a great deal of repetition has not appeared too aversive to the bright student. Whereas the rate at which the successes of the program should occur is unknown, experience certainly suggests very strongly that the rate should be high even if not continuous as suggested here. Teaching programs developed in the fashion indicated by Skinner demonstrate one thing clearly: logical and systematic subject matter has its own intrinsic motivating properties if presented correctly. Extraneous motivators may be avoided, such as items that say "Now isn't that fun, children? Let's try another one." We have found absolutely no need for pep-talks by teachers, comical visual aids, or any of the other gadgets intended to amuse and motivate the student. Their progress through the program seems to be its own reward and is sufficient to sustain progress.

One of the problems that arises in leading the student through a program concerns the decision about where to place statements that summarize the characteristics of a set of problems. For example, in teaching the student about conditioning, one might introduce the law of type S conditioning, teach the student the law, and then proceed to give examples of this type of conditioning. On the other hand, one might give many examples and gradually lead the student into a summary statement of the nature of these examples, this summary statement being the law of type S conditioning. If a rule or generalization is given before the examples, it may in some way aid the student in working the exemplary problems. On the other hand, a series of examples may themselves be sufficient to lead the student to induce the general principle. This second procedure has been generally favored in writing programs since, frequently, the summarizing would not be understood by the student until after he has seen the examples. Experience dictates the wisdom of following the series of examples with a generalizing statement in the program rather than leaving the generalization to the student's verbal descriptions. In teaching algebra, we have observed the following anomaly. The student is first given a rule that when multiplication and addition appear in the same problem, the multiplication should be done first. He is then given a number of problems of graduated difficulty. Two students, at the end of these examples, were able to work the problems quite well, but when asked whether addition or multiplication should occur first they stated "addition." Their verbal descriptions are in variance with their behavior in this case, and may at some later date lead them into difficulty. It is suggested that a program lead the student into an active verbal generalization of his behavior at the appropriate points.

D. Sequence of Self-Instruction, Expository Teaching, and Simulated Field Experience

Here we are concerned with just where self-instructional materials should occur within the total teaching program. Many types of subject matter are taught in part by expository materials, such as lectures, movies, etc. Furthermore, one kind of field experience or another may be utilized. A question is whether the self-instructional material should occur in some particular time relation to these other experiences.

Field experience, and to a lesser extent exposition, necessarily contain a large amount of irrelevancies: responses not required by the student and stimuli not requiring a specific response. These materials generally exist to give context to the required behavior. Priority of machine instruction may have the advantage of readying the student to separate relevant and irrelevant behavior in actual field experience. Furthermore, if we take the view that exposition alone is a poor instructional procedure, it may serve usefully one or more of the following functions:

a) Orientation to the task to be learned.

b) Displaying a context for the learned behavior in a fashion more complete than possible by self-instructional procedures.

c) Summarizing what has been learned.

d) Displaying the relevancy of what has been learned to other behaviors already in the student's repertory. (This is what we sometimes call "inspiring," demonstrating implications, and changing attitudes of students.)

That is to say exposition can display a situational context for the student's existing repertory. All of these functions but the first one would normally occur after self-instructional procedures. Orienting exposition would occur prior to self-instruction or may be built into the teaching device by the use of panels. Informally, we have discovered that in the teaching of psychology students who were given text assignments and demonstrations after self-instruction were:

a) More capable of observing relevant data, were more critical readers and observers.

b) Read text assignments more rapidly and with less confusion.

c) Evidenced less tendency to report that the text was "too hard" and other unfavorable attitudes towards the subject matter.

Furthermore, in some military training experiments, performed by the writer with A. J. Riopelle, it was discovered that retention of certain materials learned in the classroom was greater if the field experience followed rather than preceded the classroom instruction. It would be foolish to suggest at this time how text materials can be coordinated with machine instruction. However, we have discovered very little in the textbook that cannot be programmed. A general recommendation is that, for most instruction, the student begin immediately with the self-instructional materials and that expository materials and simulated field experience follow some unit of the self-instruction program.

The procedures outlined above by no means exhaust the principles of programming. These principles were designed to be particularly relevant to continuous discourse material. Furthermore, they were outlined with visual presentations by teaching machine or similar device in mind. Suggestions for procedures to use auditory and other kinds of visual presentation of problems has been neglected here simply because of lack of experience with these kinds of operations. The writer has chosen Skinner's* kind of programming as a baseline for analysis because of its plausibility, its demonstrated success, as well as the fact that more experience has been had with this type of programming than any other. Arguments over whether Skinner's kind of procedures are preferable to other procedures is a waste of time. In the first place, these principles are only a crude approximation to what must be optimal principles, and in the second place, any research program on principles of programmed teaching must begin somewhere. The place to begin is with the kind of program that has some degree of plausibility, the possibilities of control presentation, and some repertory of experience in its use.

III. THE USE OF MACHINES TO CONVEY TEACHING PROGRAMS

It cannot be overemphasized that teaching devices are merely conveyances for teaching programs. The characteristics of the teaching machine will, thus, be determined by what you want the teaching material to do, and the nature of the teaching material should not be dictated by the design of any particular machine. Although the existing teaching programs have been devised for one or another kind of device, it is quite conceivable that future programs will require various devices for their presentation. For an example, instruction in language may require a device for presenting visual materials as well as one for presenting auditory stimuli.

---

* Skinner cannot bear the responsibility for principles as stated here, although many of them represent the writer's interpretations of Skinner's procedures. These procedures are contrasted with the so-called "intrinsic programming" of N. Crowder, where the program is written in such a way that the material presented to the student depends upon the nature of his previous errors.

There appear to be two general philosophies that form a basis for the selection of a device for research purposes. From one position it is argued that the research program should begin with a very simple device that will perform the functions of presentation and confirmation. This device could be a mask or folder, the Porter machine, or the University of Georgia model. This philosophy is directed by the belief that research emphasis should be placed upon the development of programs and that modifications in the machine will naturally follow from what functions it is discovered that the program should perform. The advantages of this position are primarily those of economy. Such automatic teaching devices do not represent much of an investment and when research indicates modification, little is lost. These devices take ordinary typewritten or mimeographed paper, and the materials can be easily and inexpensively changed. For research purposes, the feasibility of additional devices can be examined merely by using them in conjunction with the simple device. Limitations in the position are marked by the necessity of the development of new devices when changes in the teaching program indicate new means of conveyance.

A program is best written and revised by having plentiful and continuous feed-back from typical students. The success and failure of these students will inform the writer of the value of the items in the program. Freedom to make rapid changes in the program while it is undergoing development demands a flexible means of presenting the materials. If the writer places items on cards and presents them by hand to a student, he will not only be free to make experimental changes with the maximum of ease, he will also be placed in a position in which it is virtually impossible to ignore the student. First hand and intensive observation of students actually learning will take the place of concern about mechanical breakdowns and group supervision. It is time enough to devise a machine when one knows precisely what he wants the program to do.

A somewhat opposite philosophy of machine selection would argue that a research machine should have the capability of performing any of the operations that the conveyance of a teaching program might require at one time or another. When one begins to examine the variety of things a teaching device might be called upon to do, it becomes obvious that such a device would be considerably complicated and expensive. This marks its major disadvantage. The investment in such a machine may represent money spent for functions that are later found to be quite unnecessary. A real question, I suppose, can be reduced to deciding whether one wants to build a machine that anticipates all possible functions of a teaching device or to introduce these functions as required. The two positions are not necessarily incompatible, however, since one can use both kinds of devices, the more complicated one providing information about how the more basic device can be modified. It must be remembered that here we are talking about research devices rather than schoolroom models. Skinner's machines promise to be smoothly operating devices designed for use in the classroom. They seem, however, to have some drawbacks as research machines. The major drawback lies in the inability to change around the materials easily, and it is constant revision that will take place in the research laboratory. His machines have been devised on the basis that they are to be used in a classroom with a program that has been demonstrated to teach efficiently. They are intended primarily as a teaching, not research, instruments.

At any given stage of development, the selection of an automatic teaching device will depend upon two things: a) the analysis of what a machine can do, and b) a decision about what it is one wants the machine to do. With reference to the first of these problems, there are four general classes of operations that could be required of a machine. The device can be used to 1) present material to a student in the form of information or problems; 2) receive and store a student's response; 3) confirm the correctness or incorrectness of the response; and 4) make decisions about what parts of the material are to be selected for presentation. Variations in design of automatic teaching devices will represent variations in the ways in which these operations are carried out.

A. Presentation of Materials

1. Modality of Presentation. Most automatic teaching devices are committed to visual presentations, although coordination with other visual aids and with auditory presentations is quite feasible. Such coordination has been rather extensively experimented with by those who have worked with automatic devices in the instruction of languages. The simplest kind of visual presenter can rely upon a certain degree of cross-modality transfer and present visual stimuli calling for handwritten responses when in fact, the training program is designed to teach the student to make oral responses to auditory stimuli. An example of this can be found in the Service Representative Training Program where customers' statements are presented visually in the frame of the machine and the student is required to make a handwritten answer. Later the student will be required to respond to the same stimuli when they occur over the telephone, and his response will have to be in an oral form. To what degree, and with what kinds of materials, this cross-modality transfer can be relied upon is not known. The problem does point to an area of research, however. There is no need to commit a classroom to extensive auditory devices unless their need has been demonstrated.

2. The Mechanics of Presentation. Decisions must be made about how the material is to be advanced through the machine. With the simple masking device, or folder, the student pushes the materials under the folder until the problem is exposed in the frame. The Porter machine or the University of Georgia model allow the student to advance the materials by pushing a lever which engages a gear advancing the paper on which the materials are written. Skinner's most recent machine is electrical and the materials are advanced when the student pushes a button. Ease of advancement is paid for by money, and one must decide how much ease of advancement is worth in dollars and cents.

3. The Type of Material Required. Materials can be produced on paper, on cards, on tapes, etc. For a research machine it is clear that the materials should be easy to rewrite and replace. Devices such as a card sort and the simpler teaching machines make this quite feasible. As we have said, Skinner's latest machine has the materials reproduced in such a fashion that it is difficult to make changes in the materials. This is a very severe restriction on a research device. The University of Georgia model has an advantage over Porter's machine in that more materials can be written on less paper, and has the disadvantage of this convenience costing money.

4. Accessibility of Materials to Students. One of the advantages of Skinner's newest machine is that the materials are in a cartridge such that the students can load the machine without ever handling the materials themselves. The use of a folder has no such convenience, since the material must be given to the student, giving him the opportunity to read it before it goes into the device, to mix it up, or to lose it. The way out of this is to have a teacher closely supervise the use of the materials. Porter's device has a similar disadvantage. The University of Georgia model has materials locked into the machine so that, although the students are required to handle the materials to some extent, they cannot remove them or read them while they are in the machine. The decision as to whether to allow these materials to be accessible to students will necessarily be based upon experiences with this accessibility. More complicated machines can be administrative devices giving automatic supervision of the use of the materials. Whether this is necessary or worth the expense will have to be seen.

B. Reception of the Student's Response

1. Reception of the Responses by Manual Operation. A student can present his response to automatic teaching devices in two general ways. He can recognize a correct answer when it is produced in multiple choice form and inform the machine by pushing a button or moving a slider. Secondly, he could reconstruct the answer in a non-multiple choice situation by manipulating typewriter keys, sliders or buttons. Manual operation is obviously the most desirable procedure in the case of multiple-choice problems. The use of manual operation for inserting a reconstructed response may be very convenient where the number of elements from which a response can be constructed is small, as in the case of arithmetic. Skinner's arithmetic machine is of this nature.

2. Handwritten Reception. Most of the present automatic teaching devices receive a handwritten response. This response is covered with glass, or plexi-glass, so that the subject cannot change it. The handwritten response can be made on a separate adding machine tape as in the case of Skinner's machine or the University of Georgia model, or it can be made immediately on the materials on which the problems are written, as in the case of Porter's machine. The second procedure makes it impossible to re-use materials. Some handwriting detection devices have been constructed at the Bell Telephone Laboratories and it is possible that in the future such devices may be quite useful in allowing the machine to detect characteristics of the handwritten response. The use of a folder or mask requires that the response be either written on the problem materials, in which case the materials cannot be re-used, or the response be written on the sheet of note paper. This last procedure allows the student to see his former answers, and this might be disadvantageous. Just how disadvantageous is a subject for research. Sometimes it is not feasible for the subject to write his answers in the machine. In the case of service-representative training, the response to the problem may require the student to write on a form of some kind, such as a contact memorandum or a toll-ticket. A procedure is suggested that would allow such responses to be coordinated with the automatic training device. The subject would write his response on the contact memorandum, or whatever other form that may be used, and would refer to the answer box of the machine which would tell him which of a number of contact memoranda at the top of the device he should look at. He could then look at a properly filled out contact memorandum and match

it with his own. The contact memoranda, or whatever other forms, probably should be in some scrambled order so that this student could not easily discover the correct form before he filled out his own. This procedure may be generalized to any kinds of problems that call for written answers on the part of the student which are not easily reporducible in the answer window of the device. These may be figures that he must draw, diagrams of switching circuits, the filling out of forms, and the like.

3. The Reception of Vocal Responses. Experimentation with automatic procedures for receiving vocal responses and decoding them in one fashion or another, is indicated particularly in those programs that are designed to train students to use their voice. An example is the Service Representative Training Program which aims to train the student to make telephone contacts. Of course, vocal reception can be done with any number of standard recording devices. The problem lies in coordinating this reception with the sensing and confirming by the machine. Some experimentation with this has been done in the case of the use of automatic training devices in teaching languages. Generally the confirmation is performed by having the machine give back what is considered a correct vocal response for the student's comparison. This confirmatory technique may be perfectly adequate for many kinds of automatic training device procedures.

4. Allowing the Student to Change an Answer. It seems clear that any automatic training device should allow the student to change an answer before receiving confirmation, except in those cases where the student's first answer is desired for one reason or another. The Skinner, Porter, and Georgia models all prevent the student from changing an answer once information has been received, since the answer is then placed under glass where the student can see but not erase the handwritten response. In the folder or masking device it is quite easy for the student to erase the answer and write in the correct answer. Informal observation has revealed very little of this answer changing with the use of the folder. However, the answer the student provides before confirmation supplies the primary data for understanding how a program is to be revised. It is assumed that the prevention of such changes would be an integral part of the design of the automatic training device. It is a research question as to whether the student would receive value from being able to write in the correct answer once he has discovered that his answer is wrong. If such a course is decided upon, obvious changes in the design of the existing automatic training devices would be indicated.

C. Confirmation

The term confirmation has been suggested by Porter to replace the term reinforcement, since the latter term begs the question as to what value there is in the exposure of a correct answer.

1. Three Ways to Confirm the Correctness of the Response. In the Skinner, Porter and University of Georgia models of automatic training devices, the correct answer is displayed to the student after the machine has received and stored his handwritten response. This confirmation is independent of the kind of response the student makes. He receives the correct answer whether his answer was correct or not. A second procedure would allow the student to have confirmation of his response only if it were correct. Skinner's arithmetic machine is of this variety. If the response

is wrong, the student is informed that it is wrong simply because he cannot go on to the next problem, but nothing about the correct response is made known to him. He must now insert a new response into the machine. The multiple-choice devices, when strictly adhered to, can utilize this second type of confirmation. The second type of confirmation requires the machine to sense the difference between right and wrong, and naturally has the disadvantage of added complexity and expense in the case where the responses are drawn from a virtually unlimited verbal repertory. It is easy to arrange in the case where the answer repertory is of a limited variety as in multiple-choice and number answers. Whether the second procedure has any advantages over the first is a researchable question. Our programs have been written in such a way that the probability of the student making a correct response is so high that the question of what is to be done with a wrong response is minimized. Whether the machine should allow the student to see the correct answer after he has made a wrong answer will depend upon whether the program is written with a high degree of probability of a wrong answer.

The writer has devised an automatic training device that utilizes a mediating step which reduces the unlimited repertory of verbal responses to a limited class of responses that can be more easily sensed by the machine. In this procedure the student writes in his answer on the answer tape and pushes a button which briefly opens a device that resembles an addressograph. On this "addressograph" are written twelve multiple-choice answers, one of which is correct. The student then matches his answer with these twelve answers, and if one is a match, he sets a selector button indicating the number of that answer. The machine will then inform him whether he is right or wrong. The addressograph type device is open only for a brief period of time. This procedure allows the student to tell immediately if he is right but reduces the probability of his being informed of the correct response in the case where his answer is incorrect. Such devices are complicated and should not be invested in until a) it is determined whether the program should be written to insure a certain number of errors on the part of the student, and b) whether it makes any difference if he sees the right answer after he has made an incorrect response.

3. Giving Additional Information with the Confirmation. The answer window of an automatic training device could, in addition to exposing the correct answer, provide additional information concerning the answer. It could do this whether it exposed the correct answer or not. This information would be of two varieties. It could explain why his response is wrong or give a further hint that would lead him to make the correct response on the next attempt. Following Skinner's scheme for writing programs, the procedure seems to be rather unnecessary since the introduction of such information is usually produced by augmenting in the program itself.

D. Selection of Program Materials

We are herewith concerned with the automatic training device as an instrument for sensing characteristics of the subjects' answers in such a way that the presentation of the program will be partly contingent upon the history of the students' answers. In a general way, a machine may do such things as a) dropping out problems to which correct responses are made and retaining those to which wrong responses are made, b) deciding that the student has induced the principle involved in a certain set of materials and thus allowing him to skip over further examples of the principle, and so on. The selective characteristics of the training device are designed to take care of individual differences among the students.

1. Discrimination Between Right and Wrong Responses. Skinner's college machine is so arranged that if a student marks his answer as being correct, the machine automatically drops his problem out of the program. Those problems which were not marked correct will be re-presented on another occasion. The advantages of this procedure are as yet unknown, and the technique is paid for by the added complication and expense of the machine. The Porter machine and the University of Georgia model lack this arrangement and result in simpler, less expensive devices. The drop-out procedure can be argued against on several grounds. In the first place, the present set of programming principles would lead one to write a program minimizing the number of wrong answers and indicating that the program needs revision where wrong answers occur with any frequency. These principles imply that the perfect program is one in which a student will not make wrong answers and, this being the case, the drop-out procedure is superfluous. Of course even in the "perfect" program students will make wrong answers because of inattentive reading and the like, but it is presumed that the program must be sufficiently redundant to allow for such wrong answers without materially interfering with the course of instruction. The second argument against the drop-out procedure is that most classroom subject-matter has a certain continuity and interdependence. If you pull one item out of the program and re-present it, it will often result in a problem that will appear meaningless because it is out of context. A third argument against the drop-out procedure has to do with the situation where a student may miss one or two items in the early part of the development of a concept, but learn enough about the process to get the next ten or fifteen items correct. To have to repeat the first item may have aversive effects and may be non-instrumental in the course of learning. The fourth argument against the drop-out procedure has to do with the advisability of having sufficient items representing a given concept so that rather than repeat one item, the program will have other items of a similar nature. We do not want the student to memorize the answer to a given factoring problem and so there is no reason to repeat a factoring problem once it is missed, so long as the program has enough factoring problems to insure his inducing the principle. If he has missed a factoring problem, it is better that he get a new kind of factoring problem to avoid the danger of his memorizing an answer to a particular one. This rational type of analysis indicates that the drop-out procedure would be of more value where specific paired-associate learning was contained in the program. Obviously these questions will not be resolved by rational argument and, in the final analysis, will be answered in the laboratory. The arguments have enough plausibility, however, to suggest the use of the simpler machine for research purposes until such time as the value of the drop-out procedure is understood.

2. Retracing. It is conceivable that there would be a value in an automatic training device that could sense the fact that a student was not making proper progress through a program. Such a sensing method would detect the student who was performing sufficiently poorly to justify a repetition of some earlier parts of the program, and would allow the successful student to continue. A mechanism for this detection could be built into a device by arranging for the apparatus to move back to some earlier part of the program after, say, a given proportion of failures in any given

part of the course. Or the detection could be made with respect to the failures on certain kinds of items, leading the device to go back to material that previously led up to the items failed. Another procedure would be to have the whole topic of the program repeated after a given number of failures. In teaching psychology with a folder, the writer followed some such procedure when he required the students to review a topic upon missing as many as 10 percent of the items in that topic. More elegant devices for detection of failures may make use of key items placed in the program to determine how much and what kind of material should be retraced. It would be well to avoid investment in such complicated machinery until research displays its necessity. It appears that retracing will arise when a device detects failures in the solution of problems. The philosophy of fixed programming indicates that programs should be written to minimize these failures, and that programs will anticipate the difficulty that would ordinarily lead to retracing. An alternative to retracing, possible with the simpler automatic teaching devices, is the use of programs designed for students operating at different levels of capacity.

3. Sidetracking. Whereas retracing refers to the detection of failures in the progress of the student, sidetracking would require a device that would detect unusual performance on the part of the student. An imaginary example of a sidetracking device would be one which would automatically shift to additional and more complicated material on a given subject when the student had a string of successes, and his responses were of short latency. In this imaginary example the machine is deciding that a student, on the basis of his rapid progress with the program, is prepared to learn something about the subject-matter that slower students would not be taught. The easiest form of this type of detection would, of course, be with reference to sequences of successful answers and the latency of response. A machine that would detect something about the quality of answers is conceivable though it would certainly be difficult to arrange. Naturally, a human could be used for the purpose of this type of detection. Sidetracking can be combined with retracing in the case of the student who is making insufficient progress through a program, indicating not only his need to go back to the earlier parts of the program, but the need to give him additional information in the earlier part.

4. Skipping. Another possibly useful device would detect the fact that a student needs no more problems of a given type. This device would then automatically skip to the next topic. For an example, if the student performed 10 successive factoring problems correctly, the device may then "conclude" that the student has induced the principles of straight-forward factoring, and would lead him immediately into the next steps of the program. On the other hand, the student who proceeds through the program either slowly and/or with a number of failures would not be automatically skipped to the next part of the program until he had reached some criterion of mastery. It must be remembered that both speed and accuracy can be utilized to detect the level of mastery of the student. Only research will determine how we should treat differently those students who work rapidly with many errors or those who work slowly with few errors.

5. Program Selection by the Student. In the future some thought may be given to the possibility of arranging an automatic teaching device so that the student can make selections. This would allow him to repeat certain materials, get further information about the subject matter (sidetracking), or allow him to skip on to later materials. There may be some value

in student selection, as opposed to machine selection, on the grounds that a student may have forgotten previous material on which he had performed correctly, or decided that he didn't understand the material even though he could perform the problems. An imaginary device may have on it a window in which a statement may say: "You have done well with this material. If you would like to know something about the history of behaviorism before the time of Pavlov, push button No. 7. If you would like to study some more recent experiments on conditioning, push button No. 8." There is no gainsaying that such procedures produce complicated devices, and that there is a strong argument for continuing with the presently recommended, simpler devices until research indicates how this selection can be optimally arranged.

E. Crowder's Scrambled Book Technique

Norman Crowder provided us with an ingenious technique for using something on the order of a text book* as an automatic training device. Very briefly, a page in the text book will present some information about a subject-matter and lead up to a problem. On the same page it will provide a set of multiple-choice answers, one of which the student is to select. For any alternative he chooses, there is a direction to turn to some other page in the book. The pages are not arranged in any sequence, so the student does not go from Page 1 to 2 to 3. He goes to the page indicated by his answer, and if he is correct he is told so on that page, and given further information. If he makes an incorrect choice, he is directed to a page that tells him that he is incorrect, tells him why, and gives him the information that he is assumed to need to overcome his deficiency. Crowder's scrambled book allows the use of many of the techniques described above, without a mechanical device. For example, retracing can be arranged in such a manner that if the student chooses a wrong answer of a certain kind, the scrambled book writer has anticipated the kind of information that this student needs. He can then be directed to a page in the book that gives that information, and to this extent the book can be tailored to the individual student. It can be seen that sidetracking and skipping are also possible. Crowder's technique has other advantages in that it requires no mechanical devices, can use all the techniques of presentation normally used in text books, and generally can be arranged to follow the principles of programming. Unfortunately, we know even less about the value of the scrambled book as a teaching device than we know about the Skinner procedures.

Strictly on *a priori* grounds the writer has preferred the Skinner technique for the following reasons: 1) The scrambled book in its present form allows for no storage of the student's responses. He makes a choice of an answer, but what this choice is is not recorded, and naturally he can change his choice at any point. That is to say, as a research technique it has the disadvantage that the experimenter cannot describe the course of the student's learning in any reasonable fashion. Of course, these objections can be overcome by a few manipulations, although feasible ones may bring us right back to mechanical devices. It is also possible to modify Crowder's technique so that the student is not given multiple choice answers. He could be required to write in his answers and instructed to go to a page which will tell him the correct answer, and will also anticipate possible wrong answers, giving him the further instructions that those answers would indicate. 2) It can be argued that something may be lost in the learning procedure when the student is required to make a sub-vocal response or when he is not required to commit his response to writing. In a Skinner-type teaching device, a student has to write in an answer before he gets confirmation. Where this is not required, the student may merely read (perhaps without understanding) and then choose some alternative without investing effort. To the extent that this happens the scrambled book technique will take upon itself the disadvantages of textbooks. This, of course, can be avoided by the modification suggested above, where the student is required to write in an answer. 3) A third and *a priori* argument can be made against the multiple-choice technique upon which Crowder's scrambled book is dependent. This argument has been given in an earlier part of this report, and is essentially based on the possibility that there are real differences between recognition and reconstruction of answers.

---

* Crowder also has developed machines to perform the same functions as the scrambled-book. Since the writer knows so little about the machines, the present analysis is directed to the book.

REFERENCES

1. Ferster, C. B.
2. Husband, R. W. Television versus classroom for learning general psychology. *Amer. Psychol.*, 1954, *9*, 181-184.
3. Porter, D. A critical review of a portion of the literature on teaching devices. *Harv. Educ. Rev.*, 1957, *27*, 126-147.
4. Pressey, S. L. A machine for automatic teaching of drill material. *Sch.* & *Soc.*, 1927, *25*, 549-552.
5. Skinner, B. F. Science of learning and the art of teaching. *Harv. Educ. Rev.*, 1954, *24*, 86-97.
6. Skinner, B. F. The experimental analysis of behavior. *Amer. Scient.*, 1957, *45*, 343-371.
7. Wischner, G. J. and Scheier, I. H. Some thoughts on television as an educational tool. *Amer. Psychol.*, 1955, *10*, 611-614.

SHORTCUTS AND ANALOGUE DEVICES IN ATTEMPTS TO TEACH THE MATHEMATICAL TRANSFORM CONCEPTS OF FEED-BACK THEORY TO GRADUATE STUDENTS IN PSYCHOLOGY

George L. Kreezer
Washington University

I. Introduction

The present report deals with a pre-Sputnick attempt, carried out intermittently over a span of about twelve years, to find ways to raise the level of mathematical competence of a somewhat neglected section of our school population. Though guided by a clinical consideration of the psychological factors responsible for mathematical disabilities, it had its origin in an interest in using mathematics to help psychology rather than the reverse. With respect to methodology, it may be regarded as having its locus in the exploratory period of an investigation, during which one is trying out different procedures and looking for fruitful leads, rather than in the later more formal period characterized by research designs.

The population was made up of students in courses given by the writer, initially at Cornell University, but principally at Washington University, in fields such as Cybernetics, Biomechanics, and to some extent, in Physiological Psychology. The level of mathematical training of the large majority of students in these courses was approximately that represented by conventional mathematical requirements for a bachelor's degree in a Liberal Arts college, outside of areas such as Physics, in which mathematics is an accepted tool-subject. That is to say, all of the students had had the equivalent of a course in college algebra but only occasionally did students appear who had taken courses at somewhat higher levels, as in analytic geometry and calculus. The students were almost all graduate students in psychology, with a strong interest in psychological problems, but not at all convinced that mathematics, outside of the procedures of statistics, would be of any value to them.

At this point, it may be of interest to indicate briefly how the mathematical concepts we proposed to teach came to be selected as suitable in the training of psychologists. During the war, in connection with work on feed-back systems at the Radiation Laboratory and in the Servo-mechanisms Laboratory of M.I.T., I had become acquainted with some of the mathematical methods being developed for analysis and synthesis of feed-back systems. One of the general methods used, the Laplace Transform, I found of particular interest due to its possible suitability for dealing with some of the basic concepts and problems in psychology. Consequently, following the war, some additional work was undertaken to determine the applicability of transform methods to specific biological systems (8,9), and to develop more effective techniques for deriving the necessary mathematical functions from empirical data on response (10,13,14,19).

II. Special Merits of Laplace Transform as a Mathematical Tool in Psychology

Following is a list of what seem to me to be some of the chief reasons why the Laplace transform is well qualified as a tool for theoretical work in psychology and certain other biological fields.

(1) It provides, first of all, a way of representing in mathematical form the "stimulus-organism-response" schema which is so widely used in psychology for description of response and attempts at prediction.

Of major interest is the fact that through the use of the Laplace transform one can derive a mathematical function to represent the response-system alone, without including any terms standing for the input signals or the characteristics of the response. When differential equations are used to represent systems, their stimuli and their responses, it is in general not possible to segregate the properties of the system from the characteristics of the input and the response. It is the possibility of setting up separate functions to stand for the system, the stimulus or input, and for the response, which makes the Laplace transform, or equivalent techniques, suitable for handling the stimulus-organism-response schema of psychology.

(2) The Laplace transform approach provides, secondly, an extremely simple procedure for predicting the response from a knowledge of the properties of the input and the properties of the response-system. A function representing the response can be determined simply by multiplying the function representing the input by the function representing the response-system.

It would be difficult to conceive of a simpler way mathematically to predict response. Possibly the only operation that might be regarded as simpler would be the use of addition in place of multiplication. Addition can in fact be substituted for multiplication in the present instance through the use of logarithms. Then the logarithm of the response function becomes equal to the log of the stimulus function plus the log of the system-function. Wherever multiplication is the basic operation, addition is an alternative procedure through the use of logarithms.

It seems to me that students in psychology, for whom prediction of response is such a central problem, should have the opportunity of becoming acquainted with the possibility of predicting response through the use of these relatively simple techniques. By being exposed in some measure to methods developed in fields outside of psychology, they may possibly avoid the provincialism of believing that the prediction of response of response-systems can come about only through the relatively blind procedures of verbal or statistical correlation. Through acquaintance with the existence of such powerful techniques as the Laplace transform, their levels of aspiration may possibly be raised.

(3) The Laplace transform approach also makes it possible to get around an obstacle that has prevented psychology from following the lead of physics in the widespread use of differential equations as a mathematical tool. This obstacle consists in the fact that psychologists, or biologists, do not generally have sufficient information about the components and structure of any given system, and of the laws governing processes inside the system, to write differential equations for the system and to thus use such equations as a basis for prediction. Through the Laplace transform approach, coupled with the use of non-dimensional parameters, it becomes feasible to derive system-functions on the basis of observations of the form of the response of the total system to specified inputs, without the necessity of full information about internal structures and processes in the system.

(4) Finally, the application of the Laplace transform permits appropriate mathematical and scientific problems to be translated into a form that can be solved by the simple procedures of elementary algebra. This is a circumstance that has led to such widespread use of the Laplace transform in engineering, and which opens the way to equivalent widespread use by psychologists. It is a circumstance that should appeal particularly to graduate students in psychology.

It is not possible in the space allotted here to illustrate these points. Evidence concerning some of them can be found in a summary prepared for the Radiation Laboratory Series on the concepts and methods of Servo Theory (12), in books on the application of the Laplace transform to engineering problems (e.g., 3,7), and in reports of research relating to attempts to use transform and feed-back concepts in psychology and biology (1,5,8,9,13,15,17,20,21,22).

III. Teaching Procedures

We are now in a position to review the procedures adopted for introducing students with relatively meagre preparation to the Laplace transform and its applications in psychology. A consideration of the mathematical background of the students participating, together with informal discussions designed to determine attitudes to mathematics, suggested that the following three conditions were significant in the design of a suitable course:

(a) The belief of many students that mathematics, outside of statistics, has little to offer psychology.

(b) Inadequacies in mathematical preparation, in terms of traditional standards.

(c) Feelings of insecurity of students in relation to work in mathematics.

A. Concerning the Value of Mathematics in Psychology. Discussions with students indicated a quite common belief that training in mathematics would be of little value in psychology, apart from the use of statistical procedures. Many students indicated, in fact, that they had not taken more mathematics since they saw no particular relevance to the fields in which they were interested. Whether this is a genuine reason or a rationalization is hard to determine. Nevertheless, it seemed desirable in fashioning the present course to present mathematical concepts and methods in relation to their possible value in attacking important psychological problems. The following two problems were proposed as of general importance in psychology: (1) finding more effective ways of predicting the response of systems such as are dealt with in psychology, and (2) developing techniques for representation of any given response-system, in a way that would make it possible to predict the response of the system to specified inputs. Mathematical concepts that were subsequently introduced were considered in relation to achieving these goals, and a basis was thus provided for evaluating their significance for psychology.

B. Inadequacies of Mathematical Preparation of Students in Psychology. If we consider typical fields in which the Laplace transform is widely used, such as in servo-system design or in electrical engineering, and ask what are minimal mathematical course-requirements before a student in these fields has the opportunity to become acquainted with the Laplace transform, we are likely to find a sequence such as trigonometry, algebra, analytic geometry, differential calculus, integral calculus, and differential equations. In contrast, we find that the typical graduate student in psychology has not gone beyond algebra.

The student in psychology obviously lags far behind the student in these other fields in his level of mathematical preparation. Is it necessary, in order for him to find out how to use the transform techniques in question, that the missing courses all be made up? To attempt to impose this requirement would be unrealistic, since very few students would undertake to complete the set of missing courses in the absence of any conviction of their importance; it is even conceivable that some of them would not have the aptitude to do so.

(1) Shortcuts. Is any alternative approach possible? One special feature of the Laplace transform, mentioned in the list given of its merits as a technique for psychologists, suggests a way out. The application of the Laplace transform to problems formulated in terms of derivatives or integrals translates them into relatively simple problems in algebra. The student with adequate preparation in algebra would therefore be competent to solve problems to which the Laplace transformation has been applied, if he could be equipped with the techniques and understanding necessary for making the transformations.

A consideration of the time spent in the special courses in analytic geometry, differential and integral calculus and differential equations, indicates that much of the time is spent either on concepts or in practice not essential in preparation for use of the Laplace transform in psychology. This circumstance suggested the possibility of fashioning a course for students in psychology in which all preparatory material other than that relevant to our special purpose was eliminated. On the basis of such careful selection of preparatory mathematical materials, it was found possible to design a one or two semester course aimed at equipping students, with no preparation beyond algebra, to understand and use the Laplace transform in problems of interest to psychologists.

(2) Textbook. In carrying out this program, it would obviously be helpful, as in any course, to have a textbook available that the student could use himself to reinforce and supplement material presented in the classroom, one involving the same objectives and sequence of topics. No such textbook was available. As a partial substitute, however, use was made of the summary of servo-theory prepared for publication in the Radiation Laboratory Series to which reference was made above (12).

This account fitted in with our special teaching objectives in several respects: (1) in the use made, throughout this summary, of the Laplace transform as a basis for representation of response-systems and for prediction and correction of system performance; (2) in the assumption of relatively little mathematical background on the part of the reader (although the assumed level was somewhat higher than we would specify for a specially designed textbook); and (3) in the use of an outline for classifying topics in the field that corresponds to questions that a psychologist might raise in attempting to appraise and rehabilitate an individual, psychologically.

These questions were: (1) how does a given system perform? (2) how does its performance deviate from accepted standards or norms? and (3) how can the system be corrected or modified so that deviations in performance will be eliminated or minimized? These questions were proposed as a basis for a survey of methods available for dealing with non-living servo-systems. But they might be used equally well as a basis for corrective work in technologies involving the living organism, such as medicine, education, and clinical psychology.

(3) Concrete Models and Teaching Aids. In the fields of elementary teaching, particularly in remedial work, the importance of concrete teaching aids to balance the abstractness of mathematical concepts is generally recognized (4). Our attempts to teach graduate students in psychology the concepts under discussion here suggested a parallel need at graduate levels. For this reason, a number of physical models which had been developed to help in the derivation of the system-functions of biological response-systems were utilized.

These models consisted of: (1) an electrical system-function simulator, in which separate passive units could be used to represent the factors in the non-dimensional form of any system-function; (2) a static slide-rule type of model by means of which any reasonable system-function could be represented by a spatial pattern of sliders in a slide-rule framework; and (3) a Fourier Transform model, described originally by Born, that shows how the mathematical definition of the Fourier Transform, a special case of the Laplace Transform, can be physically realized in the form of a composite electronic-optical device (2,10,16,19).

These models were all of the analogue type and permitted students to see and feel, in concrete structures, the same type of organization that was involved in our basic mathematical concepts and functions. These devices could also be used to predict and compute system-response, by processes analogous to those carried out by means of mathematical symbols. Such analogue models seem to me of much greater value for instructional purposes than digital type computers, which although they may be capable of greater computational precision, do not place in evidence the patterning involved in mathematical concepts and functions.

C. Feelings of Insecurity in Mathematics. A third condition that we attempted to take into account, in our presentation of the course-material, was the sense of insecurity frequently felt by students in the area of mathematics. Although many of them explain their failure to take more mathematics by pointing to its lack of relevance to their dominant interests, it is apparent in many cases that there is a basic feeling of insecurity, a lack of confidence in one's ability to perform mathematical operations of any complexity, or of apprehension that any new concept or problem may be beyond one's capacities.

To counteract these attitudes and anxieties, a number of fairly obvious steps were taken, such as: (a) providing reassurance by pointing out the manner in which mathematical ideas are built up in step-wise fashion, and that an accidental lack of understanding or of adequate teaching at any point in the developmental structure might result in a lack of assurance in the use of later concepts; (b) provision for a review of basic algebraic procedures of interest here, and practice in their use to establish a level of confidence;

and (c) care in building up the necessary preliminary ideas in simple stages before some key definition or concept, such as the Laplace Transform, is presented. Thus with respect to the latter concept, which is defined as a special example of a definite integral (7, p. 104), it was necessary for the student to understand the nature of differentiation, and of the reverse process, integration, resulting in an indefinite integral; the meaning of the definite integral; and the specific procedure for integration of expressions containing an exponential function, such as is involved in any Laplace transformation. Steps were taken to relate these operations to simple processes with which the student was already familiar, and an effort made to avoid names and symbols that might frighten him until the ideas were clear. We might think of this as clinical psychology applied to the teaching of mathematics.

As a result of the combination of procedures described here, I believe that it is possible, in a one or two semester course, to introduce concepts such as the Laplace Transform to students with no preparation beyond algebra, and to get them to recognize the relevance of such concepts to the empirical and theoretical problems of psychology. Such recognition can in turn have the effect of increasing their general interest and assurance in mathematics, and make them willing to take the additional work necessary to provide a more solid foundation.

The essential point suggested is that there may be two different approaches to be considered in trying to get students to take a certain sequence or array of courses in mathematics as part of a training program. One approach is to require all basic courses to be taken first, a relatively complete job being done in teaching the material logically belonging at each stage. This procedure may work well for students who have a natural predisposition or aptitude for work in mathematics. It does not seem to work well for the others. Witness the prevalent attitude to mathematics among students outside of the physical sciences.

The second approach is overtly goal-directed. It would involve avoiding all non-essential material in the preliminary courses, for the purpose of demonstrating to the student as quickly as possible that specific mathematical methods are of direct value as a means of solving problems in which the individual is already interested. Once work in mathematics has thus acquired a positive valence, in Lewin's terminology, through the operation of the law of effect, then it will be appropriate to direct the student's attention to the need of filling in the gaps in his mathematical training involved in the initial demonstrations of its value.

This procedure may be likened to that often followed in work upon problem-solving in animals. The animal is first given the actual experience of eating in the goal- or food-box, before being placed at the starting point to find his way to the goal by processes of discovery and learning. The student without special mathematical aptitudes may require equivalent experiences. He must first be allowed to taste of what may become the desired fruit before he is ready to go through the labor of finding it.

REFERENCES

1. Alexander, R. M. *Developmental changes in capacity for dynamic motor adjustment as a function of intelligence level of subnormal children.* Doctoral dissertation, Washington Univ., 1953.
2. Born, M., R. Fürth, and R. W. Pringle, A photo-electric Fourier transformer, *Nature*, 156: 756-757, 1945.
3. Brown, G. S. and Campbell, D. P. *Principles of servomechanisms*, New York, Wiley, 1948.
4. Brueckner, L. J. and Bond, G. L. *The diagnosis and treatment of learning difficulties*, New York, Appleton, 1955.
5. Dosher, G. M. *Homeostatic motor adjustments in mental patients*, Doctoral dissertation, Washington University, 1952.
6. Ellson, D. G. The application of operational analysis to human motor behavior, *Psychol. Rev.*, 56: 9-17, 1949.
7. Gardner, M. F. and Barnes, J. L. *Transients in linear systems*, New York, Wiley, 1942.
8. Kreezer, G. L. and E. H. Kreezer, The form of the light response of luminous bacteria to a sudden increase in temperature and its analysis as a transient. *Jour. Cell. and Compar. Physiol.*, 30: 173-202, 1947.
9. Kreezer, G. L. and E. H. Kreezer, The transient response of luminous bacteria to sudden temperature increments and related experiments on drug action and protein denaturation, *Biol. Bull.*, 93: 196, 1947.
10. Kreezer, G. L. Studies on the application of the concepts of servo-theory to biological systems, *Yearbook Amer. Philos. Soc.*, 163-165, 1947.
11. Kreezer, G. L. The homeostatic systems associated with the primary drives, and their theoretical representation by means of the system-functions (transfer functions) of servo-theory, *Amer. Psychol.*, 3: 239, 1948.
12. Kreezer, G. L. Servo-theory, in *Electronic Instruments*, ed. by Greenwood, I. A., Holdam, J. V., and MacRae, D., *Radiation Laboratory Series*, 21: 225-355, New York, McGraw-Hill, 1948.
13. Kreezer, G. L. The derivation of the transfer functions of homeostatic systems from experimental response curves, *Jour. Psychol.*, 28: 487-493, 1949.
14. Kreezer, G. L. An inverse decibel log frequency method for determination of the transfer functions of psychobiological systems, *Science*, 110 (2869): 681-684, 1949.
15. Kreezer, G. L. Linear approach to biological control systems, *First Confer. on Feedback Control Systems*, *Amer. Inst. Elec. Eng.*, Atlantic City, N.J., 15-16, 1951.
16. Kreezer, G. L. Development of methods for determination of the components of biological systems on the basis of the mathematical analysis of transient response curves, *Yearbook Amer. Philos. Soc.*, 147-149, 1951.
17. Kreezer, G. L. The dynamics of total systems in psychology and biology, Meeting National Acad. of Sciences, *Science*, 116: 524-525, 1952.
18. Kreezer, G. L. Why not apply mathematical methods to man-machine relationships, Symposium on Men and Machines, *Program, AAAS*, St. Louis, 166-167, 1952.

19. Kreezer, G. L. A transfer function slide rule and a transient simulator for use in determining system functions of psychological and biological response systems, _Program AAAS_, St. Louis, 100, 1952.
20. McGuire, J. C. The effect of target velocity and the area of error-tolerance circles upon performance in a two-dimensional compensatory tracking task, _USAF WADC Technical Report_, 54-431, 1-27, 1954.
21. Quesl, W. W. _Frequency analysis of human motor responses to auditory pitch_. Doctoral dissertation, Washington University, 1955.
22. Sorello, M. M. _Time course of extrapolar potential in nerve with reference to possible artifacts_. M.A. Thesis, Washington University, 1953.

SECTION II: PROPOSALS FOR RESEARCH ON THE TEACHING OF MATHEMATICS

The research designs contained in this section were proposed by the participants at the Conference on Psychological Problems and Research Methods in Mathematics Training held at Washington University on May 1 and 2. In suggesting these designs, the participants were asked to devise proposals which would be of a feasible nature and limited in scope. A suggested frame of reference was a research proposal which could be carried out with limited resources and in a limited time period.

The proposals are directed to unanswered problems in the teaching and learning of mathematics or to problems of research methodology which are especially applicable to the field of mathematics education.

The proposals are presented in alphabetical order by author.

Research Design:

The Relation of Abstractions to the Concrete: Should Mathematics be Taught in the Context of the Student's Major Area of Interest?

Marion E. Bunch
Washington University

Problem:

To what extent is the effectiveness of teaching a function of the relation between the student's major area of interest and the area of interest dominant in classroom instruction?

Hypothesis:

Among students of comparable ability, the effectiveness of classroom instruction in advancing understanding is a function of the relation between the student's major department of interests and the area of interests dominant in the classroom work.

More specifically, understanding and remembering what is learned will be greater when classroom discussion, examples, illustrations, exercises and problems used in teaching are drawn from the field of interests which is also the dominant or primary one for the student. For example, students with strong interests in economics should learn the principles, concepts, and operations involved in a course of mathematics more readily when the examples, illustrations, exercises and problems used in teaching are drawn from the field of economics, than when they are drawn from a field in which the student has little interest or when the concepts and principles are presented with minimum dependence upon concrete examples and problems from any field.

Rationale:

Several conditions may be noted which suggest the above hypothesis. Of the following rather generally accepted statements, some are supported by a greater amount of evidence than are others:

1. In general, when new concepts and principles can be tied to interesting material and experiences, learning is facilitated.

2. Skillful teachers try to lessen the unfamiliarity of new concepts and to make abstractions meaningful and interesting by relating them to what the learner already knows and is interested in.

3. Students can be made to feel the need to know and to search more actively for information when classroom presentation of the new is directed toward already existing departments of interests.

4. Active motivation to learn the subject matter of a course, or to do well in the course for more purely personal reasons, may be aroused more readily when the presentation of material in class is directed toward or appeals to a strong interest of the student.

5. Knowledge of progress on the part of the student in understanding new material and in identifying sources of difficulty should be greater when the new is presented in the context of the interesting and familiar.

6. Student participation in class may be expected to be greater when classroom presentation, discussion, etc., are centered in the field of his primary interests, and the increased opportunity for student participation can be expected to facilitate learning.

On the other hand, it might be hypothesized, that when other things are equal, the student who has broad interests will probably be less in need of selected presentation, i.e., presentation of abstract principles and concepts from mathematics in association with particulars from a principal field of interest.

The above hypothesis assumes that students may differ enough in their primary or major fields of interests to make it worth while to match classroom presentation and discussion of material to their dominant interests as, for example, we now do with regard to level of difficulty of a course and previous preparation for it in the form of formal prerequisites.

Experimental procedure:

1. Subdivide students enrolled in the course in mathematics into three groups, say A, B, and C, on the basis of their primary area of interest.

Tentative, meaningful ways in which the subdivision might be made are as follows: (a) on the basis of the student's expressed interest in, for example, social sciences, biological sciences, physical sciences, etc., or (b) on the basis of Spranger's departments of interests as defined by the Allport-Vernon Study of Values, e.g., economic, social, and theoretical values.

2. Select three teachers for the course in mathematics, as follows: the teachers should be selected so that each teacher represents strongly one of the three primary interest areas used in establishing the sections.

3. Each instructor is to place his section of the course in the setting of the primary department of interest in accordance with the classification of his section. Thus, the three teachers will design their own exercises and problems, and prepare classroom materials choosing illustrative material, etc., from their respective areas of interest. All problems, exercises, etc., with regard to the principles and concepts in the course are to be drawn from and discussed in the context of the respective interest area appropriate to the section for which the teacher was selected.

4. The three sections of students should be comparable as regards general intelligence, scholastic ability, and previous preparation in mathematics.

3. Each section should be composed of students as follows:

   (a) Fifty percent of the students in each section should have the same primary interest as that from which the section is to be taught. (Matched interest group)

   (b) Fifty percent of the students in the class should have primary interests in an area other than the one characterizing the section and the instructor. (Different interest group).

Treatment of results:

Achievement in the course may be measured by means of a course examination at the close of the semester. All students may take the same examination, constructed according to any one of several plans, of which the following are examples: (a) a unitary examination prepared by the three teachers working cooperatively; (b) an examination composed of three separate parts each of which is a final examination for one of the individual sections; (c) an examination which makes use of a standard achievement test, if available, in the subject.

The value of matching the interests of the students with the interest area underlying classroom instruction may be determined by comparisons of the subgroups and sections on the basis of the records of the final examination. These comparisons may answer such specific questions as:

1 – Within each section, did the matched interest group make significantly greater progress or achievement than the unmatched or different interest group? An affirmative answer would support the main hypothesis.

2 – Do the three subgroups with interests matched to classroom instruction differ more from each other than the unmatched interest groups differ among themselves? An affirmative answer would indicate that not all differences between sections are to be accounted for on the basis of the instructor variable.

3 – Are the differences between sections and between subgroups such as to indicate that effectiveness of matching interests in this instance is of less significance than the instructor variable?

4 – Are differences between groups absent or of such doubtful significance statistically that serious doubt is cast on the importance of matching student interest with the area of interest emphasized in classroom instruction?

Research Design:

Achievement Determination of Rate of Progression in Learning Mathematics

C. Ray Carpenter
The Pennsylvania State University

I. The Problem Area

It is assumed that the learning of mathematics proceeds as a logical and sequential development in which the elements are so ordered as to require a high level of mastery of consecutive concepts, principles, and operations. It is also assumed that many of the failures of students in mathematics, given adequate potential abilities for success, result in part from failure to master, to a necessary level, various basic concepts, principles and operations, the understanding of which is essential to the comprehension of subsequent units. Thus, failures are accumulative and they limit achievements as well as reduce positive motivation because of an adverse success-failure ratio.

A general observation shows that of each 1,000 students who enter engineering curricula in colleges, by the end of the sophomore year about 500 will have changed their programs, been dropped or withdrawn from college. For many of these students, mathematics is the critical "hurdle" course. For others, physics becomes critical because inadequate mastery of mathematics is an additional handicap to the comprehension of this intrinsically difficult subject matter.

Typically, when students fail courses or fail to attain a necessary level of mastery, entire courses are repeated even though some units of the failed courses have been adequately learned. This procedure is inefficient in many ways. Repeating courses retards normal progression through programs of studies and prevents the students from registering for scheduled courses, thus great blocks of time are lost, and the units a student has mastered must be repeated along with those failed.

II. Hypothetical Bases of Experiment

1. Learning is the product of time and effective effort. The content of college mathematics can be effectively ordered and programmed, and student progress through these programs can be scheduled and controlled on the basis of detailed knowledge of students' achievements.

2. Repetition, reiteration and variations, in optimum amounts, are known to be variables that increase, extend (generalize) and perseverate learning. Guidance and reinforcement of learning occurs when students are given knowledge of effects or results of efforts to learn.

3. Interactions between students can be arranged so as to increase learning of academic subjects through mutual reinforcement and cooperation.

4. Learning in one course may be affected positively or negatively by interactions with the learning demands of other courses in students' programs.

5. Programmed course content and learning activities within a course interact with students' formal or informal schedules of activities.

III. *Experimental Proposal*:

It is proposed to compare two patterns of instruction within two mathematics courses—Analytic Geometry and Differential Calculus presented during two consecutive semesters or quarters. For both patterns, the course content shall be the same, and the content shall be *programmed in light consecutive units* so ordered as to emphasize the sequential dependence of successive units. Two patterns of instruction will be compared which we shall designate conventional and experimental.

The conventional instructional pattern will consist of teaching the course in multiple sections as a quarter or semester course but following the same sequence of organization and development used in the experimental pattern. The usual standards, criteria and grades will be applied. Students will be given tests periodically and the normal number of completions and failures will be expected.

This control course will have the regular weekly schedule of hours of classroom instruction, and, in addition, the students will be scheduled for *two hours of problem solving sessions*. General problems will be assigned by the instructor for solutions.

The experimental classes of the course will be taught *unit by unit*. Classes or sections of students will not proceed from unit 1 to unit 2 to unit 3, - - etc., until after the class has achieved a high level of mastery of the unit being studied and have reached a performance criterion of "A" on the *knowledge-of-results* (reinforcement) *analyses*. Work will be continued on a unit until all students have achieved the established criterion level. The two hour laboratory period will be used as follows: 1. The instructor will continue to repeat (with variations) the parts of the programmed unit of study not yet mastered to the criterion level *by the majority* of members of the class. 2. He will provide complete and systematic knowledge-of-results on *specific* accomplishments to individual students including both inabilities and weaknesses. 3. On the basis of *complete and systematic learning analyses*, students in the class will be ordered according to their standing on the unit under study. All students will be informed of their own standings or rank order and those of their classmates. The students will be assigned in pairs to work-study groups for the laboratory and, if necessary, extra time individual study. The pairing will be between the highest ranking student and the lowest ranking one, and progressively in this order for the entire class. Thus, the mathematics laboratory period, and additional necessary extra time, will be so conducted as to place major responsibility on the students for their own learning and for mutual assistance.

Progression from unit to unit will be controlled by the achievement of the students of an approximation to complete mastery - - A level.

Those who master the unit quickly and effectively must be convinced of the advantages of "overlearning" in terms of retention, generalization, later acceleration and application. The rapid rate students must also be convinced that what is known can be taught to others with advantages accruing to their own learning of mathematics. Therefore, they shall accept responsibility for assisting the less rapid rate students. Thus, factors of social reinforcement and especially cooperative work are brought to bear to motivate learning.

The slowest rate students who impede the rate of progression in the mathematics courses will not be failed. Their program of courses shall be ordered in terms of the basic importance of each course for their curricula and for general education. If they have more than tolerable difficulty in keeping pace with the class, after two or more units at any time during the course, a low order course on the importance scale will be dropped without penalty to the student. If the students' rate of achievement in mathematics continues to be too slow, one or more additional courses may be dropped until a minimum of two other basically important courses plus mathematics is the entire course load of these students. Thus, the total learning burden for the student is reduced and needed time is gained for work in mathematics.

In the experimental classes the students should be given an agreement, after screening, that no one will be failed in the course who will cooperate fully and strictly comply with scheduled requirements in the program of instruction and study. He will only "fail" when he himself decided not to cooperate fully, to violate scheduled assignments, and to withdraw from the course.

IV. Subjects and Instructors

The subjects will be 800 college students, mainly engineers and science majors, registered in Analytic Geometry and who are expected to register for Differential Calculus the next semester or quarter.

The total population will be randomly assigned into the control and experimental groups and classes or sections.

Class size and time schedules will be controlled and balanced.

Each participating instructor will teach a class section by each of the patterns outlined above (one control and one experimental class for each instructor for each course).

V. Some Details of Design with Additive Variables

Criteria: 1. Performance in final exam
2. Attrition rates

(There is a problem of matching, because of selective attrition)

C: Classroom instruction (4 hours)
E-1: Substitute: Classroom film (4 hours)
E-2: Add: Permission to drop classes
E-3: Add: Problem-solving laboratory (2 hours) run by graduate student

E-4: Add: Criterion tests on each segment plus such modifications as are needed in the problem-solving laboratory.
E-5: Add: Extra time weekly for those needing it to keep up with passing criterion tests (dropping other courses as needed).
E-6: Add: Buddy system to provide maximal social reinforcement.

VI. Tests, Measurements and Criteria

The conventionally taught sections will be given pre-tests as will the experimental class sections. Conventional tests and testing schedules will be used with control sections as a basis for conventional grading. The experimental sections will be given periodically complete systematic instruments entitled "Learning analyses." The principal purposes of these will be to supply for students knowledge of the results of efforts to learn (diagnostic analyses) and to order the students for periodic pairing in work-study groups. Both the control and experimental sections will be given very comprehensive and thorough examinations at mid-semester (or quarter) and at the end of the semester (or quarter) even though the experimental sections may not have completed the whole course on the conventional schedule.

The procedures and same examining schedule will be followed for the Calculus course as for Analytic Geometry during a following semester or quarter.

It is possible that there will be progressive acceleration of the experimental sections of students. If this hypothesis holds, then these students will show slower rates of progression at the beginning of Analytic Geometry and somewhat slower during beginning units of Calculus than will be shown by the control classes. However, it is expected that by the end of two courses, two semesters or two quarters, both groups will have covered completely the two courses.

The criteria in terms of which the relative effectiveness of the two patterns of instruction and learning will be evaluated will be the following:

1. Results of quantitative tests and comprehensive examinations.
2. Variations of rates of progression for individuals.
3. Achievements in Calculus relative to achievement in Analytical Geometry.
4. Drop-outs, failures or withdrawals from mathematics or curricula which require mathematics (if drop-outs are allowed).

   Note: Steps must be taken to prevent a high attrition rate in the control group which, if permitted, would vitiate the achievement comparisons in the experiment.

5. Drop-outs, failures or withdrawals from college.
6. Achievements in courses like Physics which are closely dependent on achievements in Analytical Geometry and Calculus.

   Note: Investigators and instructors may be well advised to conduct a pilot experiment for a semester or year to organize the courses, develop the measuring and the learning analyses instruments, and to develop procedures before the actual experiment is conducted.

Research Design:

Systematic Exploration of Intra-Personal Correlates of Gain in Mathematics

Philip H. DuBois
Washington University

The Problem:

Since the time of the Pythagoreans, mathematics has been regarded as a key skill in human achievement. The range of mastery extends from the simple counting of the moron to high-order symbolic calculations of the theoretical scientist. Individual achievement obviously varies with the culture of the society, the need felt by the individual for mathematical knowledge, the mathematical sophistication and the teaching methods of instructors, and the rewards attendant upon mathematical achievement. The purpose of this investigation is to study systematically the relationships between measurable intra-personal variables, such as knowledge, aptitudes, interest and personality characteristics and gain in mathematics on defined levels of training.

Hypotheses:

A) It is hypothesized that the following abilities are related to gain in high-level mathematics: ability to handle verbal symbols; ability to handle non-verbal symbols; perceptual skills; memory; and reasoning.

B) Furthermore, it is hypothesized that, interest and motivation being constant, there are no differences in gain in mathematics achievement correlated with sex, socio-economic status or measurable personality characteristics.

Methods:

Basically this study could be executed with any set of predictor and control variables (denoted as a, b, c...j) measured on interval scales on any group of N individuals studying any academic subject, provided that two measures in that subject are available, one reflecting attainment at an initial point in time and the other reflecting attainment at a second point in time.

The method comprises, first of all, measuring N individuals on n scales suitable for the computations of product moment correlations, that is to say, on interval scales exhibiting linear relationships one with another. (If it is observed that some relationships are non-linear, scales may be transformed so that relationships become linear. In practice, with existing psychological instruments, linear relationships tend to be the rule.)

The matrix of correlations of all n variables is obtained.

Following this, the attainment measure at the second point in time (z_2) is residualized with respect to the attainment measure at the initial point in time (z_1). This can be effected either by finding actual residual scores ($z_{2.1}$) and correlating them with the other variables, or, more conveniently, by finding the partial covariances of the type $C_{j2.1}$, in which j is any variable in the matrix other than z_2 or z_1, and dividing these covariances by $\sigma_{2.1}$. These "part" correlations (sometimes called "semi-partials") express relationships between predictor variables and control variables on the one hand and gain in subject matter on the other.

We are now in a position to test our hypotheses systematically.

Results:

We can find the proportion of the variance of the criterion predicted uniquely by each of the predictors in turn. This is expressed as the square of a higher order part correlation, as $r^2_{(2.1)(a.q,\ b.q...j.q)}$.

The criterion variance predicted by the entire matrix of outside variables may be of interest. This is also expressed as the square of a multiple part correlation $R^2_{(2.1)(a,\ b...j)}$.

Particular hypotheses may be readily investigated. For example, Hypothesis A may be considered in terms of zero order r's or in terms of contribution of each predictor to a multiple R. Hypothesis B may be considered in terms of a multiple partial r between the gain criterion on the one hand and sex, socio-economic status and measurable personality characteristic on the other, with the variance associated with the control variables, interest and motivation, removed from both the criterion and the predictors. This multiple squared may be denoted as $R^2_{(2.1q)(a.q,\ b.q...)}$. This procedure removes variance associated with initial status from the gain criterion, and removes variance associated with any number of control variables, q, from both the criterion and all predictors.

The criterion variance may then be divided into portions: a part predicted uniquely by each single predictor, additional portions predicted by teams of two predictors, by teams of three predictors, and so on. This fractionization of the criterion should lead to better understanding of what measurable intra-personal variables are actually related to gain in mathematics in specified groups of subjects under specific conditions.

Significance of Results:

Since we are dealing at final stages of our analyses with variance ratios, or, interchangeably, with product-moment r's, statistical tests of significance are readily available.

Application of Results:

Results of studies executed in this general pattern can be used to answer questions as to the selection of students for courses requiring extensive training in mathematics.

Combined with normative information on our predictors and judgments of success level on the criterion, such studies can be used to investigate what fraction of the population is capable of successfully learning mathematics at different levels of training.

Research Design:

Investigation of Differences Between
Verbal and Mathematical Conceptualization

Rosalind L. Feierabend
Washington University

A. Statement of the Problem:

1. Purposes:

To investigate the nature of the difficulties which may be experienced by certain students who are high in verbal capacities and aptitude but who do not have equal facility with mathematical materials and concepts.

2. Rationale and Implications:

Many studies in the field of mathematics education take as subjects students who are above a certain level of I.Q. on the assumption that these subjects are sufficiently endowed to handle mathematics. If the subjects are found to do relatively poorly in mathematics courses or to have a negative attitude toward the study of mathematics, some other variables are sought to explain these failings.

Over-all intelligence does not necessarily mean an equally high capacity for all fields of study. While there is correlation between verbal and mathematical aptitudes, there may also be differential forms of conceptual thinking associated with individuals of high verbal and high mathematical aptitudes. It is hypothesized that there may be interference in handling symbolic representation for persons who are highly verbal and who translate all such symbols into words. The case would be analogous to the student of a foreign language who finds it necessary to translate every word into his native tongue, in contrast to the bilingual student who has associations and concepts available in both languages.

3. Hypotheses:

a. Individuals high in verbal capacity will have verbal associations to the symbols used in mathematical representation whereas individuals high in mathematical ability will have non-verbal associations (spatial forms or other symbols).

b. Individuals high in verbal capacity have an interference effect when dealing with symbolic representation, necessitating the translation of symbols or formulae into verbal form.

c. It is possible to decrease or increase the extent of verbal or mathematical associations made to symbols by varying the nature of a prior experimentally produced set.

B. Sources and Methods:

I. Experimental Procedures:

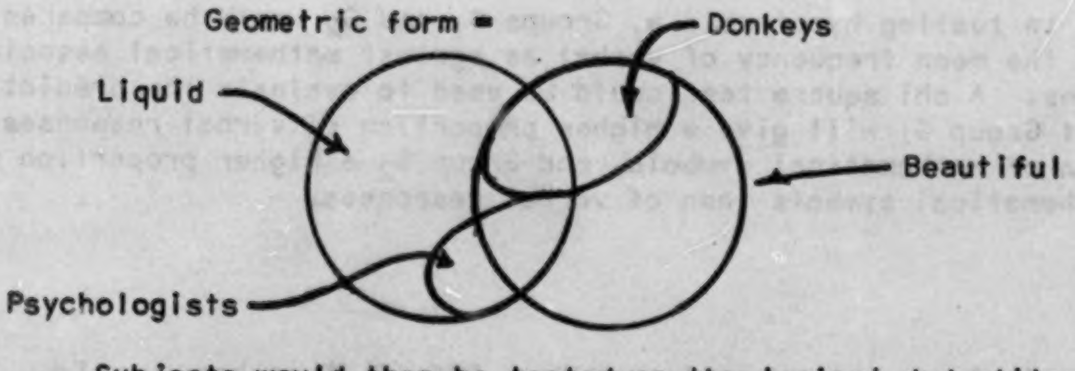
a. For hypotheses a and b two groups of subjects are needed with approximately equal mathematical background and knowledge, one group high in verbal and comparatively low in mathematical aptitude, the other group high in mathematical aptitude and comparatively low in verbal aptitude, as determined by suitable tests. Subjects could be of high school or college level.

To assess associations to mathematical symbols, both groups would be given a series of tests of the word-association type with mathematical symbols (χ, $=$, β, $+$, $\int$, etc.) interspersed among words. These tests could be both of the unrestricted and controlled association type and performed under speeded or timed conditions.

b. For hypothesis b, the same groups would be presented with a test asking them to explain a series of formulae. The term "explain" (or some similar word) is chosen because of its relative ambiguity. Explanation could be interpreted as meaning translation into verbal form or else as involving the mathematical derivation or simplification of the formulae.

Another way of determining facility for different types of conceptualization might be through a syllogisms test. With syllogisms it is possible to make a geometric representation of overlapping categories of inclusiveness or else to work from the verbal statements. Premises could be presented in both forms on the test, e.g.;

Verbal form = All donkeys are beautiful
Some donkeys are liquid
All psychologists are liquid
Some psychologists are beautiful.



Subjects would then be tested on the logical deductions derived from the premises.

Still another approach to the problem would be to present problems in mathematical reasoning to the subjects and to ask them to give, either orally or in writing, their steps to solution. Answers could be scored for verbal or mathematical content.

c. For hypothesis c, not only the two groups of subjects used in testing hypotheses a and b are required but also a heterogeneous group of subjects, that is, not divided according to relative mathematical and verbal aptitudes.

Procedure in testing this hypothesis would be somewhat akin to the Cofer studies* in which word lists were learned by subjects prior to solving the Maier two-string problem. For some groups of subjects key words for problem solution were associated on the word lists. These prior associations facilitated solution in comparison to those groups for which these key words were specifically associated with some other connotation, irrelevant to problem solution.

The three experimental groups (G_1: high verbal, lower mathematical aptitudes; G_2: high mathematical, lower verbal aptitudes; G_3: heterogeneous aptitudes) would be treated in a Before-and-After design. A control group would be comprised of half of each experimental group, selected at random.

Before-and-After tests would consist of the same tests as used for hypotheses a and b, that is, association to words and mathematical symbols, explanation of mathematical formulae, and solution of syllogisms presented in verbal and geometric form.

Experimental treatments for groups G_1, G_2 and G_3 would consist of learning lists of mathematical symbols in association with words and with other mathematical symbols and also of learning brief mathematical formulae in association with words, symbols and other formulae. One half of each group would receive lists of this type. The other half (control groups) would not receive any experimental treatment.

2. Treatment of Results:

a. In testing hypothesis a, Groups G_1 and G_2 would be compared for the mean frequency of verbal as against mathematical associations. A chi square test could be used to evaluate the prediction that Group G_1 will give a higher proportion of verbal responses than of mathematical symbols, and Group G_2 a higher proportion of mathematical symbols than of verbal responses.

---

*Cofer, C. N., in Groups, Leadership and Men, 206-17 (Guetzkow, H., Ed., Carnegie Press, Pittsburgh, Pa., 1951).

b. In testing hypothesis b, Groups G_1 and G_2 would be compared (1) for mean number of verbal explanations of formulae versus mathematical derivations or simplifications and (2) for the comparative number of correct answers on the syllogisms test to syllogisms presented in verbal and in geometric form.

Chi square could again be used to test the differences between the two groups. The prediction is that Group G_1 will give (1) a higher number of verbal explanations than of mathematical derivations or simplifications and (2) a higher proportion of correct answers to syllogisms presented in verbal than in geometric form. The reverse predictions would be made for Group G_2.

c. An analysis of convariance would be used to test hypothesis c. This analysis would show the extent of correlation between performance on the Before-and-After tests for all groups, the experimental treatment effect shown in the residuals after the regression between the Before test and the After test has been removed, and the comparison of the adjusted means of the groups on the After test.

C. Discussion:

The procedure in testing hypothesis c should be regarded as an experimental paradigm. What is of interest is not the ability to influence verbal or mathematical conceptualization by means of prior learning of word association lists. It is of interest, however, to show that type of conceptualization can be influenced. If it can, then this presumably could be done better in the course of teaching than by the learning of word lists.

This last experiment is exploratory. While it might be possible to influence performance of the heterogeneous group of subjects (G_3), it is not clear what effects it would have on Groups G_1 and G_2.

Research Design:

Ability Differences in the Learning of Concepts Governing Directed Numbers

Robert M. Gagné
Princeton University

A. Problem:

Differences in fundamental abilities appear to be prominent in the learning of mathematics, as well as in the way people use mathematical concepts. Well-established factors in human abilities are spatial, numerical, and verbal. Although there are studies which have revealed moderate to high correlations between aptitude measures and grades in mathematics, no studies have been conducted in the attempt to make specific predictions concerning the facilitation of different kinds of conceptual learning by different fundamental abilities. The possession of a high degree of spatial ability should facilitate the learning of spatial concepts; high verbal ability should facilitate the learning of verbal concepts; and high numerical ability should facilitate the learning of symbolic concepts.

The learning of concepts of addition of directed numbers may be done verbally, spatially, or symbolically. Verbal rules are perhaps the best known method, occurring in most conventional textbooks. Spatial concepts have been used with considerable success, notably in the textbook of the University of Illinois Committee on School Mathematics. Symbolic concepts can readily be designed to serve the same purpose; in one form, they might resemble some of the symbolism of Boolean algebra. Thus in this mathematical topic, the opportunity exists of relating differences in fundamental abilities to ease of learning the different types of concepts, as well as to final performance in problem solving.

B. Method:

The study requires three groups of school children, of equivalent age, probably eighth-graders. These children should have no previous knowledge of or acquaintance with algebra, or more particularly with the addition of directed numbers.

All subjects are tested on a carefully chosen factor reference battery designed to provide adequate measures of spatial, numerical, and verbal abilities. Each group is then instructed in the topic of adding directed numbers, using one of three methods (spatial concepts, symbolic concepts, or verbal concepts). Care should be taken to make the instructional materials equivalent in scope of coverage and in amount of practice provided.

When instruction has been completed, the students should be given a speeded test requiring the addition of a relatively long and varied list of directed numbers. Alternatively, two such tests could be given during instruction, making possible the use of gain scores. In any case, this type of test is used to measure degree of learning of the concepts. The test itself will, of course, constitute a practice session. Following this, students are given another test containing verbally stated problems which require the application of learned concepts to unfamiliar problems.

Correlations are obtained among (1) ability scores; (2) concept-learning scores; and (3) problem performance scores.

C. Results:

To the extent that the hypotheses underlying the study are verified, correlations between spatial ability scores and spatial concept learning scores should be significantly higher than between spatial ability scores and verbal concept learning or symbolic concept learning scores. Similar hypotheses would be tested relative to numerical ability and to verbal ability. Correlations between concept learning scores and problem performance scores, on the other hand, may reveal differences in the efficacy of different types of concepts in problem solving.

Research Design:

The Use of an Operant-Conditioning Laboratory for the Investigation of Principles of Programming Self-Instructional Materials

Thomas F. Gilbert
Harvard University

Assumptions

1. The class-room is an inefficient place to discover optimal means of programming instruction. The School Administrator's problem is to design the class-room to fit the program, not the program to fit the class-room.

2. An instructional laboratory is a place to discover the most relevant variables in programming - - not a place to test out pre-conceived variables. Statistical design requires one to specify the experimental variables prior to the experiment. The operant laboratory should provide maximal conditions for changing variables during the course of investigation.

3. Mathematics is behavior, and is subject to the laws of reinforcement and stimulus control.

4. An effective self-instructional program has two characteristics: (a) It produces the desired skills if the student goes through the program; (b) It guarantees us that he will go through the program.

5. The student is the only gauge of the efficiency of a program. For the instructional science, the student is "nature" and is never "wrong" anymore than a rock is wrong because it won't fall upwards as the experimenter wishes it to. If the student performs in a way undesired by the investigator, the program is wrong and the experiment is not over. The student fails in school, the instructor in the laboratory.

General nature of the laboratory

1. Provides for nearly complete environmental control. The experimental space can be precisely described, and does not change except by plan. The experimenter is not in the experimental space, and his instructions are minimized.

2. Utilizes the intensive study of individual students rather than statistical summaries of short behavior samples. A single subject may be investigated every day for months, or even years.

3. Provides for the complete automation of experimental programming. Every change produced by the experimenter can be explicitly stated.

4. Uses a system of continuous measurement of any behavior in question.

5. Places emphasis on control rather than prediction.

An exemplary laboratory

1. The experimental space consists of a sound-proof room, six feet square and eight feet high. The room is barren except for a chair and a panel which is attached to one wall. The room can be viewed from the outside through the ventilation system. There is a house light in the ceiling.

2. The panel has a unit appearance. However, it actually consists of an arrangement of many interchangeable parts. A typical unit may have (a) a problem screen onto which may be projected problems; (b) a shutter to cover or reveal the problem solution; (c) a reinforcement tray into which money, candy, cigarettes, etc. may fall; (d) one or more manipulanda which may be operated to produce a presentation of a problem; (e) an answer reception device, such as a keyboard for constructing the answer to an algebra problem, (f) an additional stimulus window for the presentation of such things as prompts, hints, text, or other problems; (g) a speaker for auditory presentation. All parts may be substituted with some other device, e.g., the problem window can easily be replaced with a television screen, the answer reception device with a tape for written answers, etc. A problem selector arm in the rear can be pre-programmed to select from several stacks of photo-slides depending upon the behavior of the student.

3. The panel is sufficiently flexible to allow for any arrangement that the experimenter desires. For an example, one can investigate the reinforcement value of the problems by requiring the student to operate the manipulandum in order to produce the problems. The problem presentations may be programmed on schedules just as we program the presentation of food in the pigeon or rat operant-conditioning devices.

4. Cumulative recorders provide data about such things as how the student works for problems, the rate of problem solution, the rate at which any answer is produced, the rate at which hints are produced, the rate at which the student works for hints, etc.

Scope of the laboratory

1. It may be used to model any one of the various kinds of "teaching machines" currently used.

2. Not only may we investigate the reinforcement value of steps in the program, but we may study how that value is dependent upon the correct answers: that is, we may discover the frequency at which a student requires information about the correctness of his behavior before he will extinguish his problem producing behavior (get discouraged?); or we may investigate how often he must be right in order to keep working (is the best program one in which he is always right?)

3. We may investigate the comparative reinforcement value of, say, algebra problems when the student is allowed to choose between presenting these problems or working for other things such as money, television lectures, self-instructional materials in other subject-matter, etc. We may also study the conditions necessary to increase the comparative reinforcement value of algebra or arithmetic materials (e.g. we might determine

whether frequent success is sufficient to maintain the learning behavior whether the problems are made "interesting" by gimmicks or not).

4. We may investigate the changes in the reinforcement value of the problems as experience accrues. We may investigate the satiation value of any particular method of programming, as well as compare that value with the satiation produced by competing reinforcers.

5. We can allow the student to vary the amount of "hint" he needs by making it possible for him to produce hints in a second window. It is possible to make the device a "confidence trainer" by penalizing him when he produces more hints than he needs (penalty being arranged by a time-out where nothing works).

6. Various program arrangements may be pre-arranged, or it is possible to allow the student to produce variations in the program to suit his tastes. For example, we may allow him to backtrack through the program, push buttons to make the program more or less difficult, and so on. Student can be permitted to "cheat" experimentally - - that is, to look at the answer before he makes it himself; and the conditions under which he does this can be studied.

7. The device may be used to investigate means of programming "thinking" behavior — that is, conditioning the student to shift quickly his basis for classifying discriminative stimuli when one means of classifying no longer occasions reinforcement.

8. The laboratory may be used to investigate the therapeutic value of self-instruction. For an example, it may be used to teach juvenile delinquents algebra or electronics in place of psychotherapy. Some say the delinquent is that way because of repeated failure - - particularly academic failure. It may be used to instruct "psychotic" children, with attention to its educational and therapeutic value.

In each of these classes of problems, experimental paradigms already exist in the literature of operant conditioning as guides to the actual technical procedures required. For example, Skinner has demonstrated how "mediational" behavior can be brought out into the open and investigated as any other operant - - it need not be inferred when you can require the organism to perform such behaviors publicly.

Research Design:

Placement to Maximize Efficiency of Two Methods of Teaching Algebra

Goldine C. Gleser
University of Cincinnati Medical School
and University of Illinois

Purpose: To determine whether one of two methods of instruction in beginning algebra is superior to the other, and/or whether their relative value depends on certain determinable characteristics of the students being taught which could serve as a means of assigning students to one or the other method of instruction.

The technique presently used to compare different teaching approaches is that of ascertaining differences in average performance of students taught by such methods. While it is agreed that the optimum method of teaching depends on the type of student, there have been few if any studies which have attempted to deal with this matter directly. Where placement is done it has been on the basis of intelligence or prior achievement in the subject to be taught. However, general predictors of this sort probably correlate about equally well with performance under various teaching methods and thus are not likely to differentiate those students who would do better under one or the other type of instruction. Certain personality traits, cognitive approaches, or the like might better serve as the basis of placement if they are found to be positively correlated with success under one method of instruction and to have a near zero or even negative correlation with the other.* The present research proposal is a methodological attempt to combine the experimental approach of the educator with the differential correlational approach of the tester. The teaching methods and test battery specified are intended only as illustrations.

In this study, two general methods of teaching beginning algebra will be considered. Method A consists of explicitly teaching generalizations such as use of cancellation in division, presenting the general form for obtaining the product of the sum and difference of two numbers, the square of the sum of two numbers, the solution of a quadratic, etc. Method B consists of guiding the student into discovering such generalizations for himself. One of these methods may result in superior achievement for a majority of students, but whether or not this is so it is likely that an underlying ability or trait can be found which is differentially related to the results obtained by the two methods of treatment. Thus, several measures of individual differences will be investigated as possible bases for determining those students who would benefit most from each type of teaching.

---

* Cronbach and Gleser. Psychological Tests and Personnel Decisions. Univ. of Ill. Press 1957, pp. 64-77.

Method: Students in beginning algebra in a large metropolitan high school will be assigned at random to one of six classes. Three teachers will be employed, each teaching two classes, using Method A for one class and Method B for the other. At the beginning of the year all students will be given a battery of possible predictor tests such as numerical achievement, numerical reasoning, symbolic reasoning, tolerance of ambiguity, closure, the Edwards Personal Preference Schedule, etc.

At the end of the first semester, a uniform achievement test will be administered to all 6 classes. This test would be carefully designed to incorporate the aims of the first course of instruction in algebra and would be developed on a prior sample of students, using item analysis to establish scales, ascertain reliability, and obtain a spread of difficulty such that sufficient discrimination would be possible at all levels of achievement.

Analysis of data: An analysis of variance would be performed to determine if there were significant differences in achievement between methods, between teachers, and for any particular teacher-method combination. The results of this analysis would modify subsequent analysis of data and overall conclusions. If mean performance under the two teaching methods differs significantly, it is possible that placement would still yield further improvement in performance, although less than half of the students would be taught by the method with the lower average. However, if only interaction variance was significant, placement would have to be supplemented with knowledge of teacher qualifications which favor success with a particular teaching method.

The data for students under the same teaching method would be pooled (removing mean difference, if between teacher's variance is significant). Correlations would be run between each of the tests given at the beginning of the year and the performance test. The slopes of the regression lines relating performance to predictor test under the two methods of teaching would be compared. The test (or combination of tests) yielding the greatest difference in slope would be selected as the criterion for placement. The point of intersection of the regression lines would determine the cutoff point for placement in one group or the other.

Results would be cross-validated on additional classes. If sufficient students are available this can be done simultaneously with the validation study.

Research Design:

The Effect of Verbalization of Discoveries on Transfer

Gertrude Hendrix
UICSM Mathetics Project
University of Illinois

A. Statement of the Problem:

(1) Purposes of the investigation: To find out what would have been the effect on transfer power if the subjects in the recent experiment by Haslerud and Meyers* had been required to state each decoding principle discovered.

(2)(a) Rationale of the problem: The Haslerud and Meyers study revealed strong evidence that one operates with greater transfer power from generalizations discovered than from those given to him in explicit verbal form with an accompanying illustration. No verbalization of the discoveries was required; instead the subject applied each discovery by translating a sentence into the discovered code.

(b) Implications of findings to the area of mathematics training: The effect, if any, of injecting an intervening verbalization step before the transfer test should be significant for planning instruction and writing instructional materials in mathematics.

(3) Formal statement of hypotheses:

(a) Formulating a sentence which expresses a generalization is a different and separate process from achieving a nonverbal awareness of the same generalization; the formulation (or verbalizing) process does not increase transfer power; and such formulation, necessary as it is for testing, recording, and communicating, if it be ill-timed, may actually decrease the individual's power to use the newly-discovered generalization.

(b) Apparent losses of transfer power, if any, may be due to involuntary literal interpretation of the subject's own incorrect verbalization of his own discovery.

---

*G. M. Haslerud and Shirley Meyers, "The Transfer Value of Given and Individually Derived Principles." Journal of Educational Psychology, Volume 49, pp. 293-298. December, 1953.

B. Sources and Methods:

(1) Experimental Procedures: For material and procedures communicate with Professor G. M. Haslerud, Department of Psychology, University of New Hampshire. (Possibly some one in his department would want to carry out the study.) These materials would be modified so that:

(a) The codes would be given for none of the problems.

(b) In half of the exercises the discovered principle would be stated by the subject. (However, each group should work his half of the test not requiring verbalization before being given the other half of the exercises.)

A further comparison of results from correctly verbalized principles with results from those incorrectly verbalized should be carried out. A statement of a generalization should be considered incorrect if a counter-example can be found.

(c) Subjects would be chosen as in the Haslerud and Meyers experiment.

(2) Treatment of Results: Data would be recorded to provide comparison of performance on problems in the transfer test from generalizations discovered but not verbalized, with performance from generalizations discovered and verbalized. Treatment of results should follow the procedure in the Haslerud and Meyers study.

Research Design:

Empirical Methods for Organizing Mathematics Curricula

Winton H. Manning
Washington University

A. Introduction:

The organization of topics within a course or the determination of the optimum sequence of courses in any area of learning is recognized as a fundamental educational problem. When faced with an apparently heterogeneous array of topics, the teacher must seek to impose an order or arrangement which he feels will maximally benefit the learner. Sometimes the order is dictated by the levels of complexity of the various topics; that is, the student must have mastered topic "a" before he can go on to topic "b" and so forth. In other instances the topics are related, but on an equivalent level of complexity, so that either "a" or "b" may be studied earlier.

What is apparent in most instances is that the order or sequence of learning is largely determined on the basis of professional judgment. Although professional judgment is not inherently unsatisfactory, it would be preferable to structure a curriculum on the basis of the behavior of the learners themselves.

Recently Guttman (1954) has proposed a theory for investigating the order relationships among variates. Termed radex theory, the method seeks to determine from a pattern of intercorrelations, the underlying order structure of the constituent variates. Two different ideas are embodied in a radex. The study of variates, all of the same kind, but differing in degree of complexity is termed simplex analysis. The relations between tests of addition, subtraction, multiplication and division are examples of variates differing in degree of complexity, and their intercorrelation matrix is described as a simplex. Extensive study of simplex analysis as a means of defining order relationships among topics in a training curriculum has already been carried out by Jones (1958).

This proposal deals with a second aspect of curricular organization - the determination of the most desirable sequence for a series of topics which differ in content, but which are on the same level of complexity. This situation is more likely to occur at advanced levels of learning, that is, after fundamental skills have been mastered.

A set of variates all of the same degree of complexity, but differing in the kind of knowledge they define may not be ordered in a "most" to "least" sense. However, they may be ordered in a circular fashion, so that the topics form a closed chain in which neighbors are most closely related. A set of variates obeying such a law has been termed by Guttman as a "circumplex." In the learning situation this would imply that, for example, of six topics to be studied, one may begin with any one of them; however, having begun at one point, the sequence of the remaining topics would be determined. Schematically the situation may be envisaged as in Figure 1.

A
F B

E C
D

Any sequence of study (ABCDEF, CDEFAB, FEDCBA) may be followed, but the sequence would then be determined by the first topic introduced, for the topics are circumplicially ordered.

B. Statement of the Problem:

The purpose of this study is to explore the potential usefulness of circumplex analysis for the organization of a mathematics course.

The method provides an empirical rationale for the ordering of different topics of parallel complexity so as to provide a sequence of learning which promotes optimal comprehension.

It is hypothesized that:

(1) A set of topics of equivalent complexity may be ordered as a circumplex; and

(2) A group of students trained in the topics in question, when the sequence of topics is circumplicially ordered, will obtain higher mean scores on achievement tests administered at the conclusion of each topic than will a group of students trained under a randomly sequenced series of topics.

C. Sources and Methods:

(1) Experimental Procedures:

To test the first hypothesis we would construct an achievement test battery made up of sub-tests covering each of the several related topics. This battery would be administered to a large sample of students who have already studied the topics. The correlations of each sub-test with each other may be computed, and from this matrix an analysis of the order relationships among the variates would be undertaken. The methods for conducting such an analysis will not be discussed here. If the first hypothesis concerning the existence of the circumplex is not confirmed, the second hypothesis, of course, becomes irrelevant.

Assuming that the topics may be ordered as a circumplex, an experimental course of study would then need to be planned in which a circumplicial sequence of topics would be determined by random assignment, or by purposly avoiding the sequence indicated by the circumplex. These two courses would differ in no way except for the independent variable of sequence.

Assignment of a group of students to one of the two courses, experimental or control, would be random. If preferred, pre-tests of knowledge and intelligence measures for each group might be obtained to insure further the comparability of the subjects.

At the conclusion of study of each topic, an achievement test covering the topic in question would be administered. The same set of achievement tests would be used for both experimental and control groups. These measurements would constitute the dependent variables. Treatment of data by appropriate statistical methods would then serve to test the second hypotheses.

D. Summary:

Two hypotheses concerning the investigation of order relationships have been advanced. The first is that topics in mathematics of similar complexity, but of different content will be ordered as a circumplex. If this hypothesis is confirmed, it would then be possible to study the relative merits of a course of study in which the topics are studied in circumplicial sequence as compared with a random sequence. A suggested experimental design for testing these hypotheses was then offered.

E. References:

Guttman, L. "A new approach to factor analysis: the radex," in P. F. Lazarsfeld (editor), Mathematical Thinking in the Social Sciences, Glencoe, Illinois: The Free Press, 1954, 238-348.

Jones, Marshall B. "The uses of simplex theory in training research," American Psychologist, 1958, 13, p. 398. (Abstract)

Research Design:

The Interaction Between Student Anxiety and Teaching Approach in Learning Mathematics

Wilbert J. McKeachie
University of Michigan

A. Problem: The Interaction between Student Anxiety and Teaching Approach in Learning Mathematics.

1. Purposes of the Investigation.

One of the current controversies in the teaching of mathematics is between (1) those who feel that in order to achieve greater understanding and transfer, the teacher of mathematics should emphasize the formal structure of mathematics and (2) those who feel that the most important goal of mathematics teaching is its use and that the teacher of mathematics should therefore stress the development of skills in the mathematical operations most needed in engineering, science or other mathematics courses.

Little empirical evidence can be adduced to the support of either of these positions and the protagonist of either is likely to see the sides of the two issues as black and white – with his own position possessing all of the virtues. Little consideration seems to have been given to the possibility that each approach may have special advantages for certain types of students, while being inferior to the other approach for other students. We know, for example, that students differ in intellectual ability and would expect that a more conceptual approach would be more advantageous for students high in intelligence than for those lower in intelligence. Similarly we have evidence that a high level of anxiety interferes with complex learning while facilitating simple learning.

Hypotheses:

1. Students high in intelligence will be higher in achievement and interest (relative to students low in intelligence) when taught by a structural approach[1] to mathematics than when taught by a traditional approach.

2. Students high in anxiety will be higher in achievement and interest (relative to those low in anxiety) when taught by a traditional approach than when taught by a structural approach.

B. Sources and Methods

1. Experimental Procedures.
   (a) Manipulated variables:
      (1) Teaching approach.

Each staff member teaching freshman mathematics would be asked to choose which of the two approaches to maximize in his section. Presuming that a number would choose each, weekly meetings would be held for each group to discuss methods of maximizing the effectiveness of that approach. For example, those using the structural approach might use a text like Allendolfer and Oakley. (This method may seem inferior to a design according to which each instructor teaches one section by each approach. I'll be glad to discuss this at the meeting.)

---

[1] Structural and traditional approaches obviously need to be defined more fully than is possible in this outline.

(2) Students would choose their sections without knowing of the experiment and teachers would be assigned randomly (if possible) or in such a way as to minimize the possibility of any difference due to hour of class, location of classroom, special loading of class with engineering and science students, etc.

(b) Measurements to be obtained:
- (1) Independent variables.
    - a. Intelligence:- SQT or other college entrance test.
    - b. Anxiety:- Haber-Alpert scale of debilitating anxiety.
- (2) Dependent variable
    - a. A test of achievement given in all sections devised by a committee representing both groups of teachers. This test should include items measuring
        - (1.) skills in mathematical manipulation
        - (2.) understanding of mathematical structure and logic
        - (3.) ability to apply mathematical understanding and skills to new problems
        - (4.) items measuring other types of achievement accepted as course objectives.[2]
    - b. A test of interest and satisfaction with mathematics. This test might include
        - (1.) ratings of value of the course
        - (2.) a scale of attitude toward mathematics
        - (3.) a check list of advanced courses in mathematics which the student wishes to take
        - (4.) an indication of the student's intended field of concentration.
    - c. Pre-tests
        - (1.) A mathematics test including the simpler problems of the final and other simple items measuring the same objectives.
        - (2.) a test of interest including items (2.), (3.), and (4.) under b.

2. Treatment of results.

Gain scores on each measure would be treated by analysis of variance with teaching approaches, intelligence, and anxiety as the independent variables.

---

[2] Since the advocates of either position sometimes differ in the objectives they believe to be important as well as in the worth of their method in achieving the objectives, either group should be allowed to devise and add sub-tests of objectives they believe important.

Research Design:

Research Problems in Mathematics Teaching

Paul C. Rosenbloom
University of Minnesota
and
Minnesota State Department of Education

I. Piaget's stages in the development of the concept of number.

A. The problems discussed here are suggested by Piaget, the Child's Conception of Number, Routledge and Kegan Paul Ltd., London. According to Piaget the child passes through the following three stages:

1. The child cannot make a 1-1 correspondence, and does not accept invariance of number under rearrangement of a set.

2. The child can make a 1-1 correspondence, but judges quantity by global appearance. Correspondence is based on direct perception.

3. The child can make 1-1 correspodence, and accepts invariance under rearrangement.

He sets up a number of different experiments to study various aspects of the child's development, and finds in each case three stages roughly described as above.

We want the answers to the following questions:

1. How are the stages for the various Piaget experiments related?

2. What is the distribution of ages at various stages, the correlation with mental age, and what is the normal rate of progress through the stages?

3. Can planned experiences accelerate progress or bypass intermediate stages?

The results of Piaget and the answers to questions 1 and 2 give information on the development of the naive intuition. The answer to question 3 is fundamental to the process of education. We have no hypotheses on the first two questions. Our hypothesis on the last question is that the answer is "yes."

B. Research procedures

1. A large random sample of children of ages 4-7 to replicate Piaget's experiments. Record chronological ages, mental ages, and stages in each experiment. The children may learn from the experiments, and this will affect their later responses. Therefore randomize the order of the experiments.

2. Make longitudinal study to obtain information on development through stages.

In questions 1 and 2 use random sample of nursery school, kindergarten, and first grade children in a large school system.

3. Ask child to set one chair for each doll. If in Stage 1, place each doll in a different room in a doll house. Spread out dolls on chairs and ask whether there are more chairs or dolls. If in stages 1 or 2, ask child to predict result on returning to doll house. Also disarrange one set gradually, one element at a time, repeating questions continually. Ask child to make 1-1 correspondence outside doll house and offer reward if found correct by checking with doll house. Similar techniques can be used with some of Piaget's other experiments to make it easier for child to make the comparison on the first trial.

Pour liquid from same container into vessels of different shapes and ask child where there is more. Ask child to predict height on pouring back into graduated cylinders, and then to check predictions. Ask child to check also with a balance. Same experiment with clay and with beads.

In these experiments obtain correlation between IQ, mental age, and number of trials needed to reach Stage 3. Study effects of distributed practice, varied practice with different problems, and rewards.

Note: Problems 1 and 2 are being attacked by Drs. Templin and Willis at the University of Minnesota Institute for Child Welfare with small samples of children. They hope to study problem 3 if they can obtain additional support for next year.

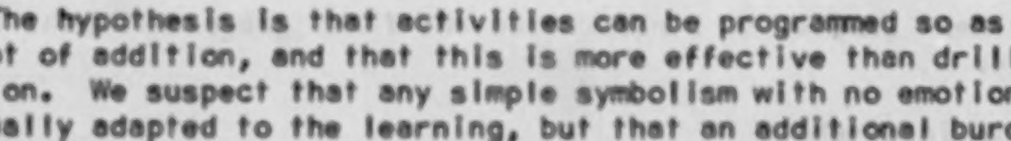
II. Concept of addition.

A. Can children be drilled in the process of addition? Is this as effective as drilling on individual combinations, as is now done in schools? What effect does the symbolism have on the child's learning? Do sensorimotor and kinaesthetic devices effectively reinforce learning?

The hypothesis is that activities can be programmed so as to teach the concept of addition, and that this is more effective than drill on each combination. We suspect that any simple symbolism with no emotional connotation is equally adapted to the learning, but that an additional burden is laid on the child when he must transfer from one symbolism to another. An understanding that the relationship is independent of the symbolism may, however, lead to greater skill in the long run and greater ability to abstract and generalize. We suspect that all reinforcements of connections have favorable effects.

B. Research procedures.

1. Take a trough separated into compartments



labeled either with numerals, letters, pictures, or other symbols. Given a set, the child is taught to place one member of each compartment beginning at the left, and to record the label of the last occupied compartment. This record may be made with flannelboard or with stencils large enough to place little demand on skill. Given two sets, the child finds the labels corresponding to each and then the labels corresponding to the union, and records the result thus:

$$C + E = H.$$

Later the child is given problems in the form

$$C + E = ?$$

and must construct the sets and then the union. He may set his own problems by picking cards at random from a deck, predicting the answer and checking the prediction. He may then make his own addition table and use it as a crutch. Later he may be able to discard it.

Find correlation between number of trials needed to learn to solve such problems and IQ, mental age, and Piaget's stages. How does this depend on choice of symbols? What part of the variance can be ascribed to the child's rote counting ability? Vary amount and distribution of practice, and study learning and retention of addition facts. Run controlled experiment on comparison with conventional method, or Skinner method, of drill on individual combinations.

2. The same technique may be used to study addition of continuous quantities such as lengths, volumes and weights. Lengths may be added with rulers and compasses, volumes with graduated cylinders, weights with a balance.

Comparisons should be made with the Stern, Cuisenaire, and Lazar materials. One should also compare effectiveness of recording results with flannelboard and with stencils.

In this experiment samples of kindergarten and first grade children, ages 5-7, should be used.

III. Use of symbols in algebra.

A. We wish to study the passage from

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \text{ and } 1 \cdot x = x \cdot 1 = x$$

to

$$(1 + x)^2 = 1 + (2 \cdot x) + x^2.$$

How do the children learn the use of abstract symbols? What aids and guidance can we give them to facilitate and accelerate this learning? How do alternative procedures compare in effectiveness? Does their effectiveness vary with the child's type of intuition?

Of course, this is typical of the fundamental problems of learning to work on an abstract level. The findings would have implications in every part of the mathematics curriculum where we wish to work with symbols and algorithms rather than meanings, yet wish to be able to interpret our work when necessary.

Some questions arise from the reports of Hadamard that some mathematicians think in terms of visual images while others in terms of abstract formalism. There are also many a priori opinions in the literature on the desirability of proceeding from the concrete to the abstract in teaching, but little experimental evidence.

B. Research procedures.

Experiments should be made with various types of preparatory programs, single and in combination, and with various follow-ups. In each case, after instruction pupils should be tested on

$$(1 + x)^3 = 1 + (3x) + (3 \cdot x^2) + x^3.$$

Correlations on performance with IQ, mental age, spatial reasoning, and abstract reasoning should be obtained, as well as time devoted to various parts of program.

1. Preparatory programs.

a. Concrete to abstract development.

Have child compute $3 + 4$, $2 \cdot (3 + 4)$, $2 \cdot 3$, $2 \cdot 4$, and verify

$$2 \cdot (3 + 4) = (2 \cdot 3) + (2 \cdot 4).$$

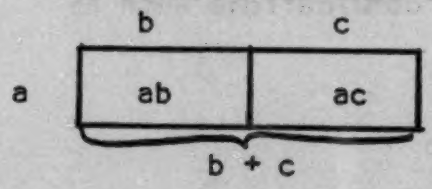
Ask him to try other choices of three numbers, and after each verification ask for general law in words. When child finally states law, give symbolic statement and ask for substitution of particular numbers for a, b, and c, and for verification.

b. Practice in substitution.

Ask child to substitute, say 3, x, and y for a, b, and c, and other permutations, also other symbols. Ask him to substitute combinations such as $(x + y)$ for a, b, or c, then substitute for several letters at a time. Ask for simplification of expressions.

c. Concrete interpretation.

Teach formula for area of rectangle. Study figure



ask for geometrical interpretation of $(1 + x)^2$.

d. **Abstract presentation.**

Present uninterpreted laws in form

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

$$(a \vee b) \vee c = a \vee b \vee c = a \vee (b \vee c)$$

$$a \vee b = b \vee a, \quad a \wedge b = b \wedge a$$

$$a \vee \alpha = a,$$

$$\alpha \wedge \alpha = \beta, \ \alpha \vee \beta = \gamma, \ \alpha \vee \gamma = \delta \ \text{etc.}$$

Show how to derive $\beta \wedge \gamma = \zeta$ thus:

$$\beta \wedge \gamma = \beta \wedge (\alpha \vee \beta) = (\beta \wedge \alpha) \vee (\beta \wedge \beta)$$

$$= \beta \vee (\beta \wedge \beta),$$

$$\beta \wedge \beta = \beta \wedge (\alpha \vee \alpha) = (\beta \wedge \alpha) \vee (\beta \wedge \alpha) = \beta \vee \beta$$

$$\beta \wedge \gamma = \beta \vee (\beta \vee \beta) = \beta \vee \beta \vee \beta$$

$$\beta \vee \beta = \beta \vee \alpha \vee \alpha = \gamma \vee \alpha = \delta$$

$$\delta \vee \beta = \delta \vee \alpha \vee \alpha = \epsilon \vee \alpha = \zeta$$

$$\beta \wedge \gamma = \zeta$$

Ask for proof of

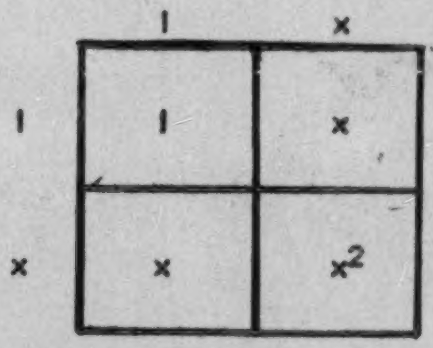
$$(\alpha \vee a) \wedge (\alpha \vee a) = \alpha \vee (\beta \wedge a) \vee (a \wedge a).$$

2. Follow-up programs.

   a. Apply $(1 + x)^2 = 1 + (2 \cdot x) + x^2$ to concrete numbers.

   b. Substitute other letters, then combinations such as x^2, $(x + y)$, etc. for x.

   c. Show geometrical interpretation



d. Show interpretation of "$\wedge$" as "$\cdot$", "$\vee$" as "+", "$\propto$" as "1", etc.

Sample should be from grades 5-9 of children without previous discussion of $(a + b)^2$.

IIIa. Use of symbols in abstraction and generalization.

A. The same purpose and rationale as III in simpler settings. Take as illustrative problem to find formula for n-th even number, and as test problems to find formulas for n-th odd number and for sum of first n odd numbers.

B. Research procedure.

1. Ask child for 1st, 2nd, 3rd, 4th, etc., even number. After each case ask for 43rd even number. How many cases before child gives right answer? Does making table help?

n	n-th even number
1	2
2	4
3	6
4	8

2. Ask for statement of rule. How can you find 100th even number, the n-th even number?

3. Ask child to find, without help, the rule for calculating the n-th odd number.

4. Ask child to make table of the sum of first n odd numbers.

n	sum of first n odd numbers
1	1
2	4
3	9

How many cases before he discovers the rule?

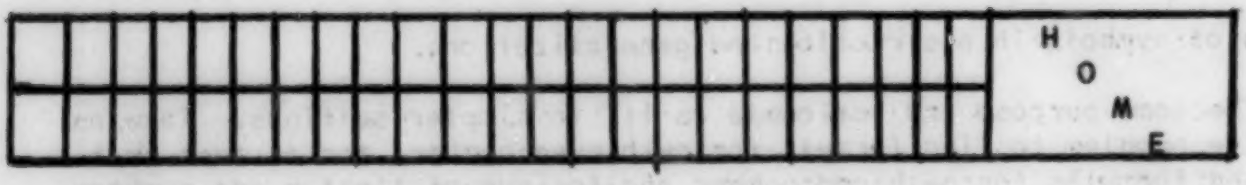
Use sample from grades 3-6. Same correlations except for spatial reasoning as before.

IV. Strategy in games and scientific method.

A. The purpose is to put the child in a highly motivated research situation and to see whether he can learn the techniques of scientific problem solving. This gives the child a chance to display his powers of inductive and deductive reasoning and his powers of generalization at various levels.

B. Research procedures.

1. Two children play on a board such as this



One counter is placed on an arbitrary square in each row. Each player in his turn moves either one of the counters at least one square, but as far as he likes, to the right. When a counter reaches home, it remains there. The player who makes the last move wins.

The children play against each other to become familiar with the game. How long does it take a child to discover the winning strategy: a winning position is one in which the counters are the same number of squares from home.

The experimenter may suggest systematic investigation by beginning with the counters near home and analyzing all possible moves.

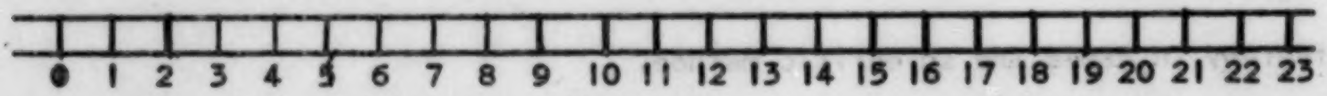
The board may be varied. One row may have the squares spaced differently. One path may be curved.

The game may be played with two piles of toothpicks. Each player in his turn takes as many as he likes from any one pile at a time. The one who takes the last one wins. The experimenter, in playing against the child, may count the toothpicks in each pile before he moves.

How many trials before the child discovers that this game is the same as the previous one?

Obtain correlation between number of trials and IQ, mental age, and abstract reasoning. Investigate transfer of learning from first version to the others. Use sample of ages 4-7.

2. A ladder is laid out with the rungs numbered.



A fireman must climb down the ladder to zero. Two children play. At first, each player in turn moves the fireman down one or two rungs. The one who brings him to zero wins.

We use the same procedure as in part 1. How long before the child discovers that the winning positions are every third rung? The game may be varied by climbing to the top. Also the maximum number of steps at a time may be k instead of 2. How long before the child makes the further generalization that the winning positions are the multiples of k + 1?

What is the effect of concentrating on $k = 2$, then on $k = 3$, etc. as opposed to having k vary at random (say depending on the toss of a die) each time? Does practice in investigating simple cases systematically in first case transfer? Take two random samples of children and teach one group the technique of scientific experimentation on the first game and the other the same technique on this game with $k = 2$. Now teach both groups this game with general k, using a pile of toothpicks, each player in his turn taking at least 1 and at most k toothpicks, and the one who takes the last toothpick wins. Which group discovers the strategy first? Partial out factors of IQ and mental age.

In problem IV, the children should be given time for free investigation between practice sessions, after instruction on research technique has been given.

V. Powers of generalization.

A. The purpose is to investigate the child's ability to generalize from empirical evidence. Does practice on such problems improve this ability? How is this ability correlated with other factors? The hypothesis is that practice improves this ability, which is transferable to dissimilar problems. Amount of transfer is probably highly correlated with intelligence.

B. Research technique.

Use sample of children in grades 4-8. Divide into 6 groups and give following 3 problems in all 6 permutations. How does number of trials until success in each problem vary with the permutation? Is a smaller number of trials needed when the problem is second or third in the series? How does amount of transfer correlate with IQ and mental age? Partial out numerical ability or arithmetical achievement. Give the child time to meditate between trials.

1. Teach the child the concept of a prime number and the sieve of Eratosthenes, and have him make a list of the primes up to 100 or 200.

The experimenter then plays a game with the child. He gives the child a number and the child must represent it as a sum of primes, using as few primes as possible. If the experimenter can find a representation using fewer primes, the child loses a point, otherwise he gains a point. He tells the child that there is a very interesting thing about the winning combinations, and that if he discovers it he will get a special prize. The winning representation is recorded each time.

How many trials before the child notices that every even number greater than 2 is the sum of two primes, and every odd number greater than 5 is the sum of at most three primes?

2. Have the child make a table of squares beginning with $0^2 = 0$. Same procedure as before with sums of squares. How long before the child discovers that every integer is the sum of no more than four squares?

Direct his attention to the representation of primes as sums of squares. Have him divide each prime by 4 and tabulate the remainders. How long before he discovers that primes of the form 4n + 1 are sums of two squares, whereas primes of the form 4n + 3 are not?

3.a. How many terms of the series

$$1 + 2 + 4 + 8 + 16 + \cdots$$

must a child add before he notices that the sum of the first n terms is less than the next term?

3.b. How many terms of the series

$$1 + 3 + 9 + 27 + 81 + \cdots$$

must he add before he discovers that the sum of the first n terms is half of the next term minus one?

Research Design:

Some Hypotheses for Psychological Research on Mathematics Training

Saul B. Sells
Texas Christian University

1. After some extensive bibliographic search, it seems that my contribution may be more profitable at this time if it deals with some hypotheses for psychological research in this area rather than a firm design for a particular study. I am still in the process of asking questions and feel that close attention to hypothesis formulation may lead to some sound research on mathematics training.

2. One basic question concerns the definition of mathematics, as a construct, and its implications for mathematics training.

Definition of mathematics. Some writers have raised questions concerning the reputedly widespread "fear" of mathematics among teachers as well as students and in the population generally. It is doubtful whether this applies to all quantitative operations inasmuch as the "fear" described seldom appears in relation to monetary calculations or other specialized quantitative operations integrally embedded in work or sports. For instance, accuracy in making change, checking costs and the like, are common. The writer has personally observed pilots compute airspeed and navigational bearings, using slide rules, who otherwise expressed dislike of mathematics. He has also observed school boys, whose arithmetic grades were poor, demonstrate considerable facility in calculating batting averages and also capability parameters of sports cars, aircraft and missiles. A related observation concerns similar relations between card playing, gambling and formal mathematics among men and women.

There appears to be a discrepancy between a number of manifestations of numerical facility, quantitative thinking, and quantitative problem solving, on the one hand, and both attitude and proficiency in formal mathematical operations, on the other. In attempting to think this through, the following questions have arisen. (a) Are these anecdotal observations valid, and to what extent? (b) Assuming their validity, as objective facts, what hypotheses appear tenable to account for them? (c) What implications do these have for training in mathematics? These are discussed briefly below.

(a) A survey of a number of age, sex and occupational population samples to assess this question might be valuable not only in opening up this line of approach, but also in demonstrating the nature of the suspected discrepancies more concretely. It would be desirable in such a survey to preserve the situational contexts of the various activities sampled, e.g. making change with money, balancing check-books, completing income-tax forms, computing rates, averages, etc.

(b) It is suspected that in formal mathematics quantitative proficiency is an end in itself whereas it is subordinated to the situational problem in the other cases. To the extent that this is true, the first hypothesis is that the situational context for mathematical operations may be an important variable, highlighting differences in "need to know," range of mathematical operations called on, terminology and formulas used in relation to problem context, and degree of overlearning of specific techniques.

A second hypothesis may involve the tendency to use formal mathematical symbols in mathematics courses, whereas in the applied situations formulas are tailored to the problem, if used at all. Related to the use of mathematical symbols is the oft-noted tendency of mathematicians to omit intermediate steps in teaching, creating problems of understanding their communications and possibly leading to what some writers have called "symbol anxiety" and "number anxiety." The abstract simplicity of symbolic notation is often difficult for learners and attention needs to be drawn to the distinction between simple in the sense of economy of explanatory dimensions and simple as the synonym of easy on a difficulty continuum. This second hypothesis may turn out to be a paraphrase of the first, but here the specific issue is abstract versus concrete.

A third hypothesis concerns the dependence of more advanced mathematics on basic operations. It is known that forgetting of all school curriculum content is rapid where the material is not used and displaced by currently more functional material. It is suggested that much of the mathematics content is forgotten and that such forgetting presents problems in later learning requiring mastery of basic material, particularly at the college level. On the other hand, operations that have functional use for particular individuals are either retained or relearned and continue to function effectively.

A fourth hypothesis concerns the use of computing aids specifically adapted to particular problems. For instance, the pilot does not use a "slide rule," but an air-speed computer. This is related to the second hypothesis, but is stated separately because the computing aid may involve an additional consideration.

The effect of this enumeration of hypotheses is to require a definition of mathematics for the purposes of investigation and training. It may be that the curriculum oriented achievement test fails to measure the range of mathematical behavior. Even in the college curriculum, mathematics may be quite different for mathematics majors and others.

(c) The problem raised in this section has implications at all levels of education. The work of Piaget on the development of numerical concepts and the literature on development of abstract abilities may require reexamination of curriculum content in terms of the concrete--abstract dimension. The problem may also reach into the approaches in teaching mathematics in business, education, physics, engineering and other departments, and in applied specialties such as navigation, book-keeping, accounting, etc. There may be a need to evaluate evidence showing that all of these specialties receiving essentially the same mathematics training in college as that given to mathematics majors.

3. From another point of view, namely prediction of success in mathematics by both ability and personality measures, the definition of mathematics is essential to the choice of criteria. Mathematicians consulted at

this University differed in the inclusion of various branches (corresponding to courses). It seems that at a minimum the formal mathematics complex should include algebra, geometry (plane and solid), trigonometry, calculus, and analytical geometry. Some mathematicians have recommended advanced calculus, as well. The classification of various operations among these rubrics is not precise. It is not known to what extent these represent homogeneous fields not what the factor structure of proficiency items is for the full range of content. Most predictive research thus far (the work at the college level is meager) has employed course grades as criteria, although instructor ratings have been used in at least one study, a "general" proficiency test.

It is proposed that proficiency tests be developed on a factored basis, without reference to course content except for developing the initial pool of items. A serious problem in this undertaking involves the location of a sufficient number of advanced students to develop adequate measures at the advanced levels. The question of adapting this approach to specialty field requirements as implied in Section 2 must be thought through.

With appropriate proficiency tests, factored ability, personality and motivational prediction studies could be designed, capitalizing extensively on results reported in the literature in developing prediction hypotheses. In such research sampling cases to control training and time interviewing since training will be a serious problem.

Research Design:

An Investigation of the Possibility of Improving Attitudes and Achievement of Students by Means of Procedures Designed to Improve Attitudes of Teachers

Donald W. Taylor
Yale University

Systematic data are available (Carey, 1955; Berry, 1958, 1959) which indicate that among college undergraduate men and women there are sizeable positive correlations between scores on scales designed to measure attitude toward analytic thinking and achievement in problem solving. Moreover, these correlations remain significant even when the effects of differences in mathematical and verbal aptitude are removed by partialling. More important, the results of one experiment (Carey, 1955) indicate that significant improvement in achievement may result from short-term group discussions designed only to improve attitude. Other data (Alpert, personal communication) show that during the school year attitudes of elementary school students tend to shift toward the attitudes of the teacher.

These data taken together suggest that: Procedures which improve the attitudes of teachers toward mathematics will result not only in improved attitudes on the part of, but also in greater improvement in performance in problem solving by, the students which they teach; the improvement in performance will remain significantly greater even when the effects of any possible concomitant increases in the teachers' knowledge of mathematics are controlled.

To test this hypothesis, instruments will need to be found or developed to measure each of the following variables:

(a) Teachers' attitude toward mathematics, or perhaps more generally toward analytic thinking
(b) Teachers' knowledge of mathematics
(c) Students' attitude
(d) Students' achievement in problem solving

Two equivalent forms would be needed for each of the four instruments.

The experiment would require the cooperation of a fairly large public school system. The procedure would involve the following steps:

(1) Administer to a large number of teachers at a given level (sixth grade?) one form of the attitude scale and one form of the test of knowledge.

(2) Administer to the students who are to enter the classes taught by these teachers, one form of the attitude scale and one form of the test of achievement in problem solving.

(3) Establish an experimental and a control group of not less than 10 teachers each, such that: (a) the two groups are comparable with respect to the mean and variance of their scores on the attitude scale and on the test of knowledge; (b) the students in the classes to be taught by the experimental group of teachers are comparable to those to be taught by the control group with respect to the mean and variance of their scores on the attitude scale and on the test of achievement; (c) insofar as the available population will permit, the teachers in both groups are chosen from those with less favorable attitudes.

(4) Over a period of at least two months, conduct at the rate of not less than once a week group and individual discussion sessions with teachers in the experimental group, these sessions being designed to improve attitudes toward mathematics. The working out of the nature of these sessions is of course a problem in itself.

(5) After not less than six months, administer to the teachers in the experimental and control groups the second form of the attitude scale and the second form of the knowledge test.

(6) At or about the same time, administer to the students in the classes taught by both groups of teachers, the second form of the attitude scale and the second form of the test of achievement in problem solving.

In every administration of an attitude scale, it should be emphasized that the data are being obtained for research purposes only and will be seen by members of the research staff and no one else. Information concerning the precise purpose of the study should be withheld from both teachers and students—a suitable "cover story" must be provided.

Employ analysis of variance to determine:

(a) whether students in the classes taught by teachers in the experimental group show relative improvement (less decline?) in attitude significantly greater than those in classes taught by teachers in the control group.

(b) whether students in the former classes show significantly greater improvement in achievement in problem solving than those in the latter classes.

If (b) yields positive results, then employ analysis of covariance to determine:

(c) whether the improvement in achievement remains significantly greater for the former than for the latter even when the effects are controlled of any possible superiority of the teachers in the experimental group in gain in relevant knowledge.

Since men and women have been shown to differ significantly in attitude toward analytic thinking (Carey, 1955; Milton, 1957; Berry, 1958) and since most elementary teachers are women, it would probably be desirable to exclude men teachers from this study. Supplementary analyses of the data would be of interest to determine:

(d) whether the effects predicted here are larger for girls than for boys.

This might be expected on the assumption that the girls would be more likely to identify with, and hence to be influenced by the attitudes of, the women teachers.

It would be of interest to include in the experiment a second control group of teachers comparable to the other two groups with respect to all of the conditions listed under Step (3) above with the single exception that this second control group would be chosen to have as high as possible (instead of low) a mean on the first form of the attitude scale. Analysis of variance would be employed to determine:

(e) whether students in the classes taught by the teachers in the second control group show significantly better attitudes at the end of the year than those in classes taught by teachers in the first control group.

(f) whether students in the former classes show significantly greater improvement in problem solving than those in the latter classes.

A study involving only the two control groups might be worthwhile, should it not be feasible to include the experimental group.

Research Design:

Experimental Evaluation of Sensory-Perceptual Aids in Learning Mathematical Concepts

James M. Vanderplas
Washington University

A. Statement of the Problem.

(1) The purpose of the investigation is to assess the differential effectiveness of various sensory-perceptual aids to learning of mathematical concepts.

(2) The rationale of the problem is based on a long-standing controversy concerning the effectiveness of sensory-perceptual aids, such as models, linkages, pictures, film-strips, and the like, in facilitating development of mathematical concepts. Some educators scorn the use of such materials on the basis that often they foster the development of "concrete" concepts and impede the development of "abstract" concepts, of prime importance in mathematics. Teaching the concepts of trigonometry, for example, by reference to angles is held by some to interfere with the use of these concepts when they are employed in problems unrelated to angles, as, for instance, analysis of wave-forms or solution of infinite series. Others hold that only by initial reference to "concrete" objects and events can the abstractions be made meaningful to the student. Building abstract concepts, such as number, thus is believed by these persons to depend on the initial establishment of relations of the symbols to objects and events experienced by the student.

Evidence exists to support both views. In some studies improvement is noted in attainment of concepts when sensory aids are employed; in other studies no improvement (or sometimes inhibition) of performance of mathematical operations is found.

It seems reasonable to suppose that for some concepts sensory aids would be useful, though for others these would not. Number skills, for example, may develop more readily if the student learns first to count by ones and then to group units into larger "wholes" (although no clear evidence supports this hypothesis). On the other hand solution to certain trigonometric problems (e.g. wave analysis) may be impeded if the student learns the functions by reference to angles. An important task, then, would be to identify whether (1) the use of a sensory aid facilitates learning of a given concept and (2) some particular aids are more useful then others for particular concepts.

The findings of such research would have obvious implications for the identification and selection of appropriate sensory-aids for teaching various concepts.

(3) The experimental hypotheses on which the study would be based are stated broadly below:

(a) Differences of unknown amounts exist with respect to the effectiveness of any particular sensory-perceptual aid (Model, graph, etc.) in facilitating the attainment of a particular mathematical concept. These differences may be zero or small as compared with other methods which do not employ such aids.

(b) The effectiveness of sensory-perceptual materials depends upon the particular experiences of the student in the area to which the aid refers, upon intelligence, and upon other unidentified factors which give rise to individual differences.

B. Method.

(1) Procedure

(a) The attainment of each mathematical concept for which sensory-perceptual aids are available would be assessed separately. The concept would be taught to a number of groups of students, but for each group a different aid would be employed. A control would also be used, in the form of a group to which the concept is taught without the use of any aid. Within each group using a particular aid, the students would be subdivided into several groups, on the basis of a measure of intelligence or past experience (e.g., social status or previous attainment, which may influence the "richness" or variety of experiences), or perhaps a measure of perceptual ability or speed. The aim in stratifying the groups is to assess differences, in effectiveness of the aids, which depend upon factors not common to all students.

Students would be randomly assigned to groups, within the restrictions noted above, to take account of individual differences.

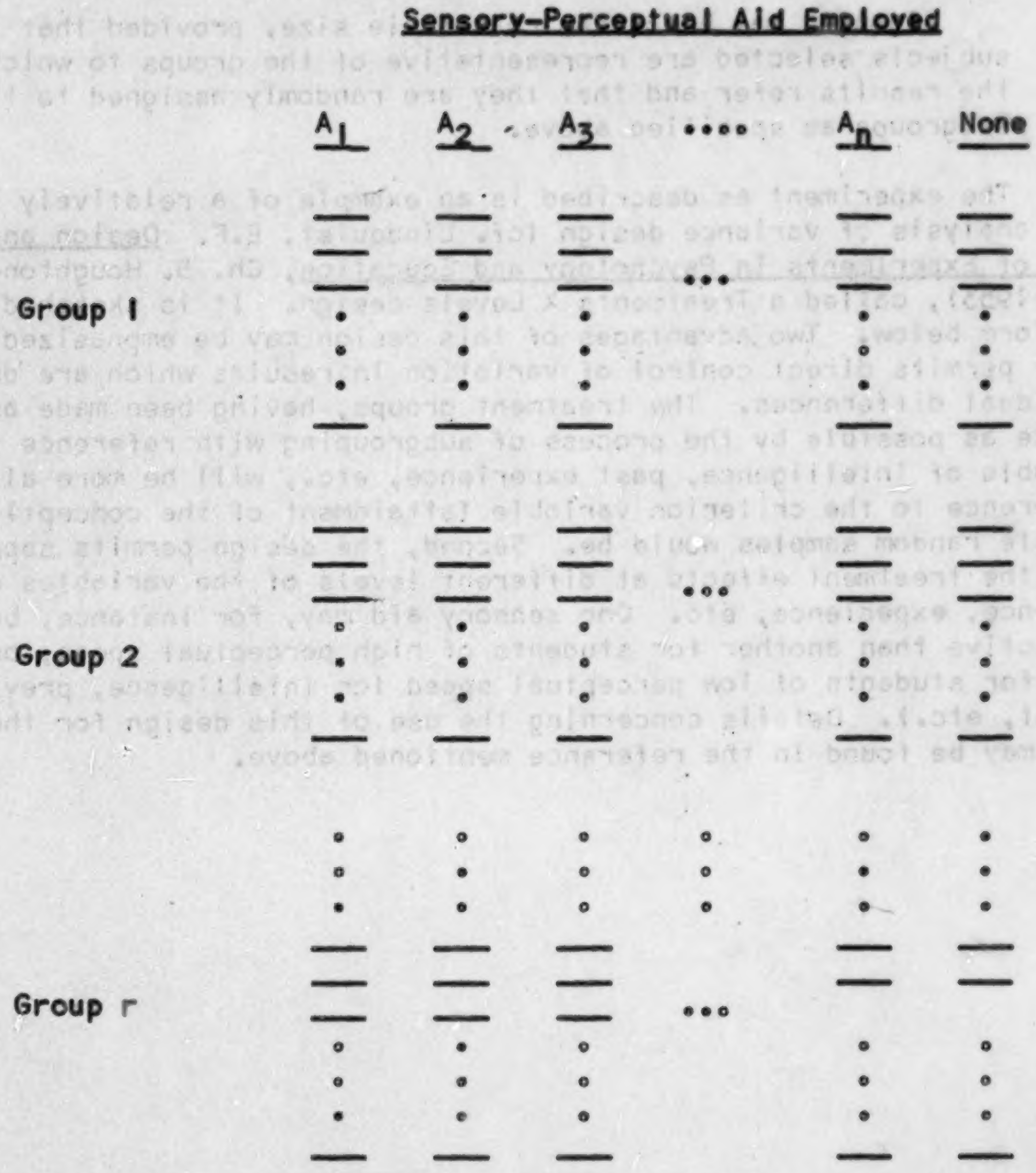
(b) An objective measure of attainment, such as a score on a test employing problems, applications, etc., requiring use of the concept, would be used. For example, a standard arithmetic achievement test could be used as a measure of attainment of numerical concepts. For problems requiring use of geometric or other higher-order concepts, tests may need to be devised. In each case, however, a single measure would be used as an attainment score (e.g., no. of problems solved or no. of errors or no. of successful applications of the concept).

(c) Subjects would be students in the usual courses of mathematics, with previous training as provided in normal curricula. We would suggest, however, that the study be carried out first in lower grades, since the effectiveness or need of sensory-perceptual aids may depend greatly on the extent to which abstract concepts are developed in the early courses. This dependence may complicate the experiment if carried on in upper grades, since it would require further subdivision of groups, according to previous experiences, in order to assess the effects of these early experiences with sensory-perceptual aids upon development of higher-order concepts.

Samples may be of any reasonable size, provided that the subjects selected are representative of the groups to which the results refer and that they are randomly assigned to the subgroups as specified above.

(2) The experiment as described is an example of a relatively standard analysis of variance design (cf. Lindquist, E.F. *Design and Analysis of Experiments in Psychology and Education*, Ch. 5. Houghton-Mifflin, 1953), called a Treatments X Levels design. It is sketched in outline form below. Two advantages of this design may be emphasized: First, it permits direct control of variation in results which are due to individual differences. The treatment groups, having been made as much alike as possible by the process of subgrouping with reference to the variable of intelligence, past experience, etc., will be more alike with reference to the criterion variable (attainment of the concept) than simple random samples would be. Second, the design permits separate study of the treatment effects at different levels of the variables of intelligence, experience, etc. One sensory aid may, for instance, be more effective than another for students of high perceptual speed, but inferior for students of low perceptual speed (or intelligence, previous attainment, etc.). Details concerning the use of this design for these purposes may be found in the reference mentioned above.

Arrangement of Concept Attainment Scores for Treatments X Levels Analysis


Sensory-Perceptual Aid Employed
A_1
A_2
A_3
....
A_n
None
Group 1
Group 2
Group r

In the above arrangement each mark (___) in the table represents the score of a different subject on the attainment measure. A_1, A_2, ... A_n refer to different sensory-perceptual aids employed in teaching the concept. Group 1, and Group 2, ... refer to the subgroups, stratified according to the variable considered appropriate (e.g., High, medium, low I.Q., or Ranges of scores on a perceptual speed test, or ranges of scores on a test of previous attainment, etc.).

SELECTED LIST OF ARTICLES CONTAINING SUGGESTED SENSORY AIDS TO THE TEACHING OF MATHEMATICS

1. Amig, Margaret C. A device for teaching locus. Math. Teach., 1941, 34, 279.
2. Anderson, G. R. Teaching the slide rule via television. Math. Teach., 1950, 43, 272-274.
3. Anderson, G. R. & Vandermeer, A. W. A comparative study of the effectiveness of lessons on the slide rule presented via television and in person. Math. Teach., 1954, 47, 323-327.
4. Anning, N. A device for teachers of trigonometry. Sch. Sci. & Math., 1925, 25, 739-740.
5. Baravalle, H. von Demonstrations of conic sections and skew curves with string models. Math. Teach., 1946, 39, 284-287.
6. Berger, E. J. (Ed.) Devices for a mathematical lab. Math. Teach., 1953, 46 (Monthly dept.)
7. Breslich, E. R. The technique of teaching secondary school mathematics. Chicago: Univ. of Chicago Press, 1930.
8. Breslich, E. R. Problems in teaching secondary school mathematics. Chicago: Univ. of Chicago Press, 1931.
9. Breslich, E. R. Visual aids in teaching. Nat. Council of Teachers of Math., 18th yr. bk., 1945, 30-41.
10. Breslich, E. R. Teaching mathematics as a system of understanding. Math. Teach., 1949, 42, 61-66.
11. Carr, Alice R. Using a "calculator" to develop basic understanding and meaning in arithmetic. Math. Teach., 1950, 43, 195-196.
12. Dent, E. C. The audio-visual handbook. (4th Ed.) Chicago: Soc. for Visual Educ., Inc., 1942.
13. Drake, R. & Johnson, D. A. Vitalizing geometry with visual aids. Math. Teach., 1940, 33, 56-59.
14. Engle, T. L. Constructing a transit as a project in geometry. Math. Teach., 1931, 24, 444-447.
15. Engle, T. L. Some suggestions for using amateur photography in mathematics courses. Sch. Sci. & Math., 1933, 33, 506-510.
16. Ewing, G. W. An optical illustration of conic shapes. Sch. Sci. & Math., 1938, 38, 276-277.
17. Fehr, H. F. The role of insight in the learning of mathematics. Math. Teach., 1954, 47, 386-392.
18. Fogelson, J. D. Use of pictures in mathematics. Chicago Sch. J., 1946, 27, 65-68.
19. Garnett, P. D. Teaching arithmetic by visual training. Opton. Wkly., 47, 237-238.
20. Gillam, E. E. Mechanical aids in teaching junior mathematics. Sch. (2nd Ed.) 1943, 32, 145-149.
21. Gorman, F. H. What laboratory equipment for elementary and high school math. Sch. Sci. & Math., 1943, 43, 335-344.
22. Hart, H. F. Tennis ball geometry. Math. Teach., 1922, 15, 239-240.
23. Hickey, May. The efficiency of certain shapes in nature and technology. Math. Teach., 1939, 32, 129-133.
24. Hoag, Jessie M. Visual aids in teaching geometry. Nat. Math. Mag., 1939, 14, 153-159.
25. Johnson, D. A. Vitalizing geometry by the use of pictures. Sch. Sci. & Math., 1938, 38, 1032-1034.

26. Kemmerling, H. A device to aid in generalizing geometric idea. *Sch. Sci. & Math.*, 1927, 27, 606-609.

27. Krathwohl, W. C. Helping mathematics with an exhibit. *Math. Teach.*, 1938, 31, 68-69.

28. Krathwohl, W. C. Teaching with color, *J. Engng. Educ.*, 1952, 43, 155-158.

29. Lynch, J. A. Relation of mathematics to the learning process. *Math. Teach.*, 1939, 32, 313-315.

30. Martung, M. L. Multi-sensory aids in the teaching of mathematics. *Sch. Rev.*, 1946, 54, 191-193.

31. Mejdak, R. J. Classroom practice in the teaching of everyday mathematics. *Math. Teach.*, 1941, 34, 368-369.

32. Metcalfe, L. S. New visual aids for mathematics. *Sec. Educ.*, 1943, 11, 11-13.

33. Miller, G. A. New aid for mathematics teachers. *Sch. & Soc.*, 1933, 38, 120-121.

34. Moore, R. V. Aids to mathematics teachers, *H. Sch.*, 1929, 7, 9-11.

35. Morres, D. L. Inverted geometry. *Math. Teach.*, 1938, 31, 78-80.

36. Neal, O. Use of figures in solving problems in algebra and geometry. *Math. Teach.*, 1940, 33, 210-212.

37. Norton, Donna. The numeral frame - a device for use in the teaching of counting, *Math. Teach.*, 1951, 44, 382-384.

38. Parker, Jean. The use of puzzles in teaching mathematics. *Math. Teach.*, 1955, 48, 218-227.

39. Parsons, Ada M. Colored crayons as an aid in teaching mathematics. *Sch. Sci. & Math.*, 1914, 14, 33-35.

40. Ramseyer, J. A. Mathematics laboratory - a device for vitalizing mathematics. *Math. Teach.*, 1935, 28, 228-233.

41. Rheims, J. J. Device for effective presentation of dial meter reading. *Sch. Sci. & Math.*, 1949, 49, 557-558.

42. Riter, H. E. Enrichment of the mathematics course. *Math. Teach.*, 1938, 31, 3-6.

43. Sample, P. A. A device for teaching the extraction of roots. *Math. Teach.*, 1947, 40, 340-344.

44. Schreiber, E. W. A study of the cultivation of space imagery in solid geometry through the use of models. *Math. Teach.*, 1923, 16, 102-111.

45. Settle, M. L. Use of visual aids in the teaching of math. *Ky. Sch. J.*, 1939, 18, 16-18.

46. Sifton, Edith. Multi-sensory aids: Some theories and a few practices, *Nat. Council of Teachers of Math.*, 18 yr. bk., 1945, 1-15.

47. Sleight, Norma. An experiment in the use of graph papers. *Math. Teach.*, 1942, 35, 84-87.

48. Smith, D. E., & Reeve, W. D. *The teaching of junior high school mathematics.* Boston: Ginn, 1927.

49. Spencer, M. C. A color scheme for figures in solid geometry. *Sch. Sci. & Math.*, 1910, 10, 549.

50. Spitzer, H. F., & Burch, R. L. Methods and material in the teaching of mathematics. *Rev. Educ. Res.*, 1948, 18, 337-349.

51. Syer, H. W. Sensory learning applied to mathematics. *Nat. Council of Teach. of Math. Yearbook*, 1945, 21, 99-155.

52. Vilsvik, B. R. New materials and equipment in the teaching of math. *Sch. Sci. & Math.*, 1939, 39, 432-442.

53. Waters, Idella. Vitalizing geometry through illustrative material. *Math. Teach.*, 1935, 28, 101-110.

54. Woodby, Lauren G. The angle mirror - a teaching device for plane geometry, *Math. Teach.*, 1954, 47, 71-72.

55. Yates, R. C. Laboratory in mathematics. *Educ. Dig.*, 1949, 15, 54-55.

56. Young, Edna H. Employing visual aids to teach locus in plane geometry. *Nat. Council of Teachers of Math.*, 18th yr bk., 1945, 109-116.

Research Design:

Patterns in Verbal Behaviors Observed in the Mathematics Classroom and in Individual Pupil Attack of Mathematics Problems

E. Muriel J. Wright
Washington University

(1) The purpose of the investigation is to make a longitudinal comparison of patterns of discursive thinking demonstrated in the discussion in mathematics classrooms and in the attack by individual pupils of mathematics problems at several stages in the development of power in logical thinking of children and adolescents.

(2) The rationale of the problem must be concerned with the importance of verbal behavior as an appropriate medium for the study of patterns in mathematical thinking, in the viewpoints that will be taken in its study, and in the consideration of stages in the child's and adolescent's development of power in logical thinking. Finally the place of the classroom as the learning situation must be noted.

First, verbal behavior: the language of the classroom or of the individual, including its mathematical refinements and its approximations by gesture, hierogliphics or drawings as an appropriate medium for the study of patterns in mathematical thinking. Langer (1) points out that "language has a form which actually requires us to string out our ideas even though their objects rest one within the other" and that this implies the limitation on language that only thought can be arranged in such an order can be spoken at all. While thought itself must include more than discursive thinking, mastery of mathematical subject matter itself implies an understanding of specific mathematical ideas which finally may be demonstrable by the discursive symbolism of language.

Secondly, understanding of the meaning of the language used in mathematical argument necessarily involves two aspects: mathematical content and psychological process. An additional factor must be considered in a learning situation--sociological attitude: for the cognitive aspects have been effected by impetus from the environment of the speaker. A convenient schema for presenting these three viewpoints to be studied in detail is that of three concentric circles: the innermost, content; the middle, process; and the outer, attitude. This schema is in agreement with Piaget's (2) affective and cognitive aspects of thinking, and with Cassirer's (3) consideration of the cognitive aspect as "thing-concepts" and "relation-concepts".

Piaget's comprehensive and intensive study of thinking in the child and in the adolescent has indicated distinct ordering of the development of the psychological processes or "relation-concepts". These may be outlined in variable detail from such examples as the three stages in development of the concept of number during ages approximately 4-7 through assurance in use of inductive and deductive reasoning as adolescence advances.

Finally, throughout the periods of many of these stages, the formal training in mathematics is carried out in the classroom--in discussion and in individual work. This provides two media for the study of the developing stages--the classroom with the teacher as an adult acting as a stimulus and for understanding present stage development of the next stage, and the individual pupil attack on problems with the actual attainment of the individual pupil demonstrated.

Assurance of the position of these stages described by Piaget and of their relationship with classroom verbal patterns would suggest modifications in classroom procedure for facilitating movement to the next stage or even later stages.

(3) The hypotheses may be outlined as follows:

i. Piaget's stages will be demonstrated in a longitudinal study of classroom discussion and individual pupil verbal problem solving. These will be related to age and ability of the subjects.

ii. The concurrent classroom discussions and individual pupils' patterns of thinking will generally be similar.

iii. The teacher patterns of verbal behavior will differ more in younger grades from those of the pupils in class discussion and in individual attack of problems than in later grades.

iv. In transitional periods (e.g. ages 9-11 approximately) the development in classroom discussion of the patterns of the next stage will precede their development in individual attack of problems.

Method

(1) Procedure

The several stages will be explicitly defined, and the necessary range of grades for study will be determined and specified by preliminary replication of Piaget's experiments. For each stage the following particular procedure will be followed:

(a) Patterns of behaviors in the areas of content, process and attitude would be determined in the classroom mathematics discussion period in general, in the study of problem solving permitting demonstration of distinctions of the given stage (use of Piaget's problems or analogs).

(b) These patterns would be studied by the Wright-Proctor Observational Schedule (refer paper, this report) with specific categories outlined under the areas of content, process and attitude as follows:
content--fundamentals, relationships, mathematical applications, other applications or historical significance, techniques
process--analyzing, synthesizing, specializing, generalizing, relevant information
attitude--curiosity, independence, receptivity
neutral--(behaviors falling outside three areas)

In addition replication of Piaget's method of analysis would occur where complete protocols were available i.e. in recording of individual's attack of problems. Desirably this should also be applied to the classroom by complete recording of discussion--ideal way sound movies, through one-way screen, class well wired for sound. Difficulties with administration in regular classroom might require use of laboratory setting for a small class in which this recording could be carried out. In the individual attack on problems tape recordings of verbal development will be analyzed.

(c) Subjects would be classrooms of mathematics i.e. teacher and pupils with pupils from these same classrooms providing subjects for the individual pupil problem solving studies. These would not be those encountered in exploratory work.

(2) Analysis

Analysis of patterns will be carried out by study of individual categories within each area by the use of the binomial distribution, and over a full area by the use of X^2.

REFERENCES

1. Langer, Susanne K., Philosophy in a new key, New American Library, New York: 1948, pp. 65, 70, 73, 79.
2. Piaget, Jean, The child's conception of number, The Humanities Press, New York: 1952.
3. Cassirer, Ernst, Substance and function, Dover Publications, New York: 1953, pp. 9, 345.

SECTION III: REVIEW OF RESEARCH ON MATHEMATICAL EDUCATION

The present section contains a survey of research in the field of mathematical education for the ten-year period 1948-1958. These studies from the psychological and educational literature are directed either to problems specifically related to mathematics training or else to problems and findings which have application to the teaching of mathematics. Following the survey is a bibliography of selected references in the field.

Among the topics covered in the survey are the comparison of teaching methods; the relation of proficiency in mathematics to variables such as thought processes, motivation, mental abilities, and personality; the problem of evaluation and measurement; and the research in this area which has been carried out in countries other than the United States.

A REVIEW OF RESEARCH ON MATHEMATICAL EDUCATION 1948-1958

Rosalind L. Feierabend
Washington University

I. INTRODUCTION

Interest in the teaching of mathematics is not only increasing but spreading to disciplines which have not previously felt this area to be one of their concerns. As one scans the literature in education and psychology journals, it is clear that while educators have been concerned with the problems of mathematics training for some years, psychologists are performing a rash of recent studies in response to a growing awareness of needs in this area.

This new awareness is not only of deficiencies in the teaching of mathematics but also of the contribution which psychology ought to make toward answering some of the questions raised in the field of teaching today. There is a trend for the professional person, the expert in the field, to take an interest in the manner in which his subject is taught, even at the elementary level. This is especially true of mathematics training. Professional mathematicians are turning to a re-examination of the teaching of grade school arithmetic and of high school mathematics. Textbook material is being re-written and new methods applied. In these experimental ventures, psychologists also have a role to play. They are in a position to suggest relevant research findings, important variables for investigation, techniques of evaluation, and directions for further research.

This review of research is compiled on the assumption that it is important to know what has been done in order to know what ought to be done. It represents an intensive survey of psychology and education journals published over a period of ten years in which two types of studies have been sought and noted: (1) studies specifically related to mathematics education and (2) studies which, while not specifically related to mathematics, are thought to contain material of interest to the teaching of mathematics or to teaching in general.

Certain general works on the subject of the teaching of mathematics (78,228)* have appeared which in some respects duplicate the functions of this article. These discuss the problems extant in mathematics education and offer comprehensive bibliographies to the reader. The focus of this review, however, is not so much problem-oriented as research-oriented. As a review of research, it excludes those references which are discussions or expositions and confines itself to attempts to explain some feature of the teaching-learning process and, more particularly, of the mathematics teaching-learning process.

---

* The numbers in parentheses throughout this section refer to the works listed in the Bibliography.

Reviews of research in the area of mathematics education also exist for the years 1952 through 1956 (29,30,174). These are comprehensive surveys of unpublished research-in-progress and are limited to research specific to the field of mathematics. Each contains an annotated bibliography giving the outlines of all studies which are discussed in the text.

Apart from these general research surveys with their comprehensive bibliographies many works exist on the psychology of teaching mathematics, the classics among which are probably those of Thorndike (239,240,241). While they have wielded a major influence on the teaching of mathematics in their day and contributed to controversy and research investigation of the relative merits of association and drill versus other methods of teaching, they are not in themselves research studies. For the same reason, all other general works on the teaching of mathematics, even those advocating new approaches (e.g. 172,225), lie outside of the scope of this discussion.

This review, then, is limited to research on the factors or variables affecting the teaching or the learning of mathematics. Methodological questions are of as much importance as research findings. Few if any concrete results seem to have emerged from the welter of studies encountered, a point which is also maintained in some of the papers presented in an earlier section of this report (McKeachie, Carpenter). If future research is to be of greater benefit to the field, it will depend upon stricter attention paid to methodology. It may well be that the limitations imposed by the traditional approach of varying but a small number of factors in a single experiment is a major handicap to research in the field of education (*cf.* Carpenter). It is also true, however, that too many of the studies performed in this area suffer from a lack of adequate control of variables, either statistical or experimental, with a consequent confounding of results.

In dividing the studies encountered into topics or categories, one may separate those studies which attempt to compare the relative merits of two teaching methods from those which seek to establish a relationship between some variable or set of variables and behavior relevant to mathematics. This first category, comparing the relative merits of diverse methods, encompasses perhaps the largest number of studies, predominantly those found in education and especially mathematics education journals. The second category covers a large number of topics, e.g., abilities, motivation, thought processes, which will be discussed in turn in the course of this review. A third category of studies is concerned exclusively with evaluation and measurement of mathematics behavior and proficiency. This is the area of test construction and validation and, again, comprises a large body of studies in both education and psychology journals.

II. COMPARISON OF METHODS

A first type of study included under this heading is not a comparison of methods of teaching arithmetic but one which advocates the use of one particular method. Among these writers there is argument in favor of teaching for "social utility" or applicability to real-life situations (260), introducing arithmetic in first grade (253), teaching for meaning and understanding (18,20,48,244,257). There is argument that meaning must be related to direct experience and, in young children (230), to manual operations and that arithmetic problems involving familiar situations are more easily solved (249,251). Finally, there is a detailed discussion (197) of a method judged successful in the teaching of division-with-a-remainder because all steps were carefully planned and directed, including pupil readiness, motivation, developmental experiences, reorganization of experience, discovery and provision for transfer.

The main body of studies are comparative in approach, that is, a comparison is made between the relative achievement of an experimental and a control group. A review by Gibb (101) discusses and analyzes eight comparative studies performed between 1940 and 1950 from the point of view of their methodology, conclusions and the types of problems investigated. These studies will not be discussed again here. One group of studies compares the "traditional" with the inductive or "discovery" method of teaching in various branches of mathematics. Sobel (223) finds the discovery method superior for the teaching of algebra concepts, but only with students of high I.Q. (100-115). Norman (179) compared three methods of teaching division facts: textbook, conventional and developmental. The latter method encouraged children to generalize. No significant differences were found between groups on an immediate recall test, but on delayed recall (two weeks later) the developmental group was superior on the test of facts which had not been taught, i.e. the transfer test. In a comparison of a developmental versus a conventional method of teaching multiplication, Fullerton (94) found the developmental method superior on both immediate and delayed tests for both recall and transfer.

Pincus (190) compared the success of drill in subtraction and addition facts where drill was based on an understanding of number relationships and where it was simply a rote performance, but found no significant differences. In another study of addition facts, Thiele (238) compared the relative effectiveness of drill versus generalization (discovery and use of relationships). He found the generalization groups clearly superior. Hagelberg (111) gave lessons to sixth grade classes in the procedure for solving verbal arithmetic problems. He found no significant differences between experimental and control (no instruction) groups on a standardized test but superior performance of experimental subjects on a test of his own devising. Illoff (129) compared the general achievement of pupils in eighth grade mathematics when material was based on students' experience and parental interest was aroused. These two conditions produced higher achievement.

Gabel (95) compared the relative effectiveness of distributed and massed class time in college freshman mathematics but found no significant differences in learning between students attending one 150 minute class period, two 75 minute class periods, three 50 minute periods and five 30 minute periods. A study by Semmelmeyer (214) advocates the relative superiority of instruction in abstract concepts (area, volume, perimeter and circumference) by using an "extensional" method based on reference to the subjects' own experience and on induction of generalizations from experience.

Other comparative studies deal with the improvement in spatial visualization resulting from the study of solid geometry (28), the benefits of a tachistoscopic presentation of multiplication facts (12), and a comparison of two methods of teaching higher decade addition, "adding by endings" and "adding by tens" (11). Also considered are the relative benefits of an arithmetic readiness program at the kindergarten level (137); the systematic arrangement of practice material in addition, subtraction, multiplication and division (168) and the effect of reading instruction upon achievement in seventh grade arithmetic (62).

A group of studies in remedial arithmetic suggests first of all the need for remedial classes (107,108), the greater benefit derived from small, "clinical" remedial classes rather than large remedial classes (16), and a comparison of methods of attack used by "good" and "poor" achievers in correcting errors (199).

Among investigations of relative teacher effectiveness in arithmetic classes, a study by Schunert (212) maintains that students' algebra achievement is related to the type of institution at which the teacher took his training and the length of his experience, more than eight years' experience being found optimal. LeBaron (148) submitted a questionnaire to teachers of elementary arithmetic to determine the extent to which they agreed with the results of research findings. He found a 50% agreement, especially with the view of teaching for meaning.

Studies also exist comparing the relative merits of methods of instruction for areas other than mathematics. Typical of such investigation are comparisons of lecture versus directed discussion methods (45), group-study versus lecture-demonstration (253), recitation, discussion and tutorial methods (106), panels, recordings and films (226), permanent versus flexible groupings of students (126) and three-versus five-hour sections (252). Results of such studies either show no significant differences between methods as measured by student achievement, or else they show that success of method varies among students differentiated according to intelligence or according to some other variable.

A final, more general type of study concerns teaching methods in broad perspective, including such variables as teacher-pupil relations and classroom "climate." Henry (122) points to the phenomenon of pupil docility: the students' need to give the answer the teacher desires. Mones (171) offers a review of educational efforts from the psychological point of view. Two studies on classroom climate and emotional adjustment (41,134) suggest that the more popular, better-adjusted pupils are the higher achievers and that "good" classroom climate promotes achievement. Some studies also exist on the relative merits of positive and negative guidance (138), and on functioning under negative conditions (167,185).

Methodological problems in studies of this type are many. The ideal paradigm of a study of comparative teaching methods involves first of all the comparison of an experimental with a control group. Usually this is not feasible and what is called a control group is actually another experimental group taught by a "traditional" method.

If it were possible, three groups should be compared: an actual control (no-instruction group) and the two or more experimental groups.

The second problem encountered is that of matching subjects for relevant variables, usually intelligence. This may be done experimentally by selection (which is difficult in practice) or by controlling statistically for I.Q. through an analysis of covariance design, or by partial correlation.

Thirdly, the variability among teachers and the teacher-classroom interaction is often high, thus tending to confound the experimental variable. Since it is frequently impossible in practice to have each teacher use both teaching methods, this is another factor which must be controlled by statistical means.

A fourth question arises over the type of test or measurement device to be used. Most often, recourse is had to some standard aptitude or achievement test despite the fact that the objectives of the instructional method are often not adequately tested by this type of measuring device. A better procedure is for the researcher to devise his own measuring instrument, calculated to tap the effect which his experimental method was originally designed to produce.

These four problems can be taken into account by investigators in this field. The remaining problem is inherent in the situation and has no immediate or ready solution. In order to measure the effect of a method, some initial measure of subject proficiency is necessary with which the final measure can be compared, that is, a gain score is determined and the relative gain between groups compared. If a true control (no-instruction) group were possible, then both a before-and-after measure might not be necessary. Performance on a final proficiency test could be compared among the three groups, taking the control group's performance as a base-line against which to compare the performances of the two experimental groups. Without this, certain methodological problems are encountered. In the first place, there may be practice effects due to the repeated administration of the same test on the same subjects. Secondly, a crude gain score does not tend to correlate with intelligence or with other variables. A discussion of the use of the residual gain score in preference to the crude gain score is contained in another part of this report (cf. DuBois) and also in DuBois & Manning (71) and Manning (162).

III. THE RELATION OF MATHEMATICS BEHAVIOR TO OTHER VARIABLES

A. Concept-Formation and Problem-Solving

Teaching in mathematics involves the learning of concepts and hence much of the literature on concept formation and problem-solving may have relevance for mathematics instruction. Heidbreder's studies (117,118) suggest that number concepts are the most difficult to attain, when compared with concepts of concrete objects and forms. This finding has come under criticism from two sources, however, (63,250), the first of which claims that Heidbreder's instances are not perceptually equivalent and the second, that order of attainment can be influenced by set.

A different aspect of concept formation and problem-solving is involved in the finding that subjects tend to become rigid or "fixed" in a certain type of solution, thus impairing their ability to solve new problems which demand another approach. A review of studies on the subject of set, rigidity and functional fixedness may be found in Taylor & McNemar (237, esp. pp. 458-465). Efforts have been made to relate this tendency to personality characteristics, to extent of prior experience with the now non-functional approach, to time-lapse between this prior experience and presentation of the new set of problems and, finally, to the extent of reward operative in the prior experience. An experiment by Youtz (263) established a relationship between number of previous successes with one approach and difficulty of shifting to a new approach. These experiments are related to Duncker's (72) classical analysis of problem-solving and the steps required to solution. A somewhat different approach to problem-solving is the analysis of the type of information given to the subjects and the relative effect of different types of information upon transfer to new problems. These studies are discussed in the section on Transfer.

A final area of problem-solving of interest to mathematics is the finding of sex differences in problem solving ability. This finding and its implications are discussed fully in another section of this report (cf. Taylor). Since these differences do not appear to be related to differences either in intelligence or in special aptitudes, it is inferred that they result either from differential motivation between men and women or from differences in personality characteristics.

There is more to be said in the area of concept formation and concept utilization than can be gleaned from a review of the literature. In the preceding section, studies (18,20,48,244,257) were quoted which advocated teaching mathematics for "meaning" and supplying meaning in terms of actual experience and even manipulatory operations. This is a recognition of the fact that mathematics is a language. The referents for the symbols in mathematics must be learned as well as the relationships among them. In the writer's opinion there is no reason why meaning has to be established in terms of manipulation, except perhaps with very young children. An experiment with 9 to 11 year olds (87) has shown that, depending upon the degree of reinforced experience between sign and referent, meaning may be established either in terms of the similar stimulus characteristics of object-referents or of the manipulations performed with a series of diverse object-referents. What is important is that meaning is established between sign and referent (184) and that the attributes which will be connected with a mathematical symbol represent an abstraction and generalization of the reinforced experience with that symbol. This experience may be in the realm of manipulation, perception or at the more abstract level of other signs (as, for example, mathematical symbols explained in words).

B. Motivation

Some general references in the area of motivation and learning include Symonds' (236) discussion of the necessity of motivating students before schoolroom learning can occur, a study by Douvan (70) in which pattern of achievement was found to be related to whether social class background of the student was middle or lower class, and an experimental investigation of persistence in school work by MacArthur (154). A study by Basescu(10) investigated the relation between intelligence and reinforcement in a card-sorting task, finding that increased reinforcement improved the performance of high I.Q. but not of low I.Q. subjects. Bouchard (27) attempted a study in group motivation with a 6th grade arithmetic class and found that while experiments in group motivation are difficult to control, experimental subjects did respond with a higher rate of achievement to the knowledge that other sixth graders were taking the same test.

1. Existing Preferences Toward Various School Subjects

A group of studies have attempted to assess student preference for the various school subjects included in the average curriculum. Chase (50) obtained preference ratings by asking pupils to rank their first three preferences among school subjects studied and then to give a rating of "like", "neither like nor dislike", and "dislike" for each school subject and a rating of "easy" or "hard". His findings generally are that the top three subjects, by preference, were reading, arithmetic and art. The second three: social studies, spelling, and music. The last four: science, health education, penmanship, and language. In a rating of teachers' preferences for subjects taught, arithmetic was at the top of the list, spelling at the bottom. Chase found no differences in preferences when students were divided according to intelligence.

In a study using the same method as Chase, Mosher(176) obtained subject preferences of 2,164 pupils in three school systems, urban, rural and mountain, grades 4-12. He found that different results obtained, depending upon whether the pupils ranked their first three preferences or rated every subject according to "like", "neither like nor dislike", and "dislike". Results showed arithmetic and mathematics to be among the first three preferences by the rank-order method in all three types of schools and for all grades. On the other hand, by the rating method, mathematics and arithmetic were among the most disliked subjects by 7th and 8th grade pupils in the mountain school and the 8th through 12th grade in the rural school.

In a study of loss of interest in high school subjects as reported by college women, Young (262) found geometry, algebra, and mathematics, in the order listed, to be subjects in which students lost interest. Loss of interest in Latin and history, however, had an even higher incidence. On the other hand, the three mathematics courses were those for which the highest percentage of subjects said they never regained interest and, after physics, were also those for which the highest percentage of subjects said they never took courses in those fields again.

Granting that there are differences in student preferences for various school subjects, if properly measured, what is the relation between preference and performance?

2. The Effect of Preference Upon Academic Achievement

A series of recent investigations have attempted to explain differences in school performance among students of equal abilities on the basis of their differential attitudes. A study by Armstrong (6) correlated scores on the Otis Quick-Scoring Mental Ability Test with average school marks for high school students. Those students with an average school grade at least one-half of a standard error below their predicted average were considered under-achievers. These were matched for age, sex, grade and curriculum with a control group (normal achievers) whose average school grades were within one-half of a standard error of their predicted average.

Subjects were then given the Kuder-Preference Record Test, Vocational and Personal, and rating scales were obtained from selected teachers on traits of cooperation, dependability, and judgment for the subjects. Results indicated that under-achievers:

(1) Chose their future occupations because of the influence of others
(2) Had future vocational goals which were not in agreement with their interests on the Kuder Preference Test
(3) Had a greater number of low scores and a smaller number of high scores on the computational part of the Kuder.

The Armstrong study is typical in methodology of the others performed in this area. Reed (200) divided high school students into three groups: under-achievers, achievers, and over-achievers by comparing the ranks of subjects in standard achievement tests with their ranks on I.Q. tests. Variables which did not differentiate between the three groups were: (1) age; (2) school attendance; (3) grade points earned; and (4) number of subjects studied. Significant differences showed that: 1) the over-achiever group had: a) nearly twice as many girls as boys; b) less general intelligence and language ability; c) was more sensitive to psychological pressures; and d) was rated higher by teachers in home and class work, class habits, attitudes toward people and toward school. 2) The under-achiever group had: a) nearly twice as many boys as girls; b) relatively high general intelligence and language ability; c) were not disturbed by psychological or social pressures; and d) were rated by teachers high in mental abilities and low in conformity to school-prescribed behavior. 3) The achiever group held an intermediate position between the other two groups on most traits.

Kurtz and Swenson (144) ranked 200 students on the basis of both intelligence and achievement. This resulted in five groups: (1) High I.Q. - High Achievement; (2) High I.Q. - Minus Achievement; (3) Medium I.Q. - Medium Achievement; (4) Low I.Q. - Plus Achievement; (5) Low I.Q. - Low Achievement. The authors then asked the students, their parents and teachers to rate their attitudes as to the following questions: (1) How the student is getting along in school; (2) How he likes school; (3) Is school achievement important; (4) Is education important. Findings were that all attitudes were more closely related to students' achievement than to their ability. The minus-achiever group, for example, ranked second in I.Q., fourth in achievement, and fourth and fifth in attitude ratings.

Georgiana (100) used undergraduates in elementary psychology courses to evaluate two motivational tests (The Iowa Picture Interpretation Test and The Survey of Study Habits and Attitudes) as predictors of college success. Using multiple regression equations, this author compared the relative predictive value of the two motivational tests as compared with: 1) rank in high school graduating class; 2) the Otis Quick-Scoring Mental Abilities Test; and 3) The Iowa Entrance Test Battery. She found that The Survey of Study Habits and Attitudes made a significant contribution to the prediction of academic success whereas the other test did not. The Survey accounted for variance present in grade-point average but absent in intelligence and achievement tests. In a similar type of study, Berdie (14) found that vocational interests are a valid predictor for the curriculum a student will follow and that, in predicting college training programs, differential interests are a better predictor than differential abilities.

McGauvran (156) used the Woodman-Durost Evaluation of Student Opinions Test to measure attitude toward school and the value of an education. When administered to students in eight high schools, the scores did not correlate significantly with high school marks, intelligence test scores or ratings for college potential. When given three replications at three different teachers' colleges, each replication yielded a slight but positive correlation with grade-point average.

A study by Shepler (216) associated achievement in high school science with preference for the subject. His conclusions are that while mental ability is the best predictor of potential accomplishment, preference becomes the best predictor when ability is held constant. Also he found a slightly higher preference for science among students of higher mental ability and among boys. In the same type of investigation, Eldridge (81) compared variations in accomplishment with subject preference and mental ability. He found that dominant subject preference varied from school to school and that when preference and accomplishment are paired, there is a statistically significant correspondence between the two.

A final study along these lines by Davis (65) intercorrelated mental ability, achievement and subject preference for eighth grade students. His findings are that relative student preference is highly related to accomplishment in mathematics, science and language, although not necessarily in social studies. He also found that preference interacts with mental ability to influence accomplishment but, contrary to Shepler, he did not find a relationship between subject preference and level of mental ability.

This group of studies maintains that preference for a subject is related to achievement once variations in mental ability are held constant. No attempt is made to discover the factors underlying preference or whether preference is the result of successful achievement rather than vice versa.

3. How to Measure Attitudes Toward School Subjects

Studies by Dutton (75,76,77) on the attitudes of high school students and of prospective teachers towards arithmetic illustrate the usual method of devising attitude scales on mathematics. Essays on the subjects' most favorable and least favorable attitudes towards arithmetic were requested from prospective teachers and from these essays a series of statements were drawn representing favorable and unfavorable attitudes. The author then constructed a Thurston-type of scale of attitude statements ranging from 1.0 (dislike) to 10.5 (like). Dutton's final study refined the statements on arithmetic gathered over a five-year period from 600 students in education into a 22-item scale constructed by the Thurstone technique.

A study by Billig (21) employed much the same procedure. Essays were first requested from 10th grade girls in a course in commercial arithmetic, expressing their feelings about arithmetic and commercial arithmetic. From these essays, a series of statements were drawn. This study differed from Dutton's studies, however, in that the statements were sorted into three piles (favorable, uncertain, antagonistic) and then sorted again into three piles according to the average grade of the author of the statement. In this three-by-three comparison, the most favorable attitude was expressed by the group with the highest average grade; the least favorable attitude by the group with the lowest grade; and the favorable and uncertain attitudes by the group with an intermediate average grade. Billig then constructed a 16-item Likert-type scale from these statements which he proceeded to use as a diagnostic device to predict, at the beginning of the year, those students who would do well and those who would do poorly in the course. He equated favorable attitude with good performance.

Bendig and Hughes (13) devised a booklet of 30 negative statements about statistics and mathematics, and presented them to 71 subjects to be rated on a five-point scale according to how well the statement reflected the respondent's feelings towards a course in introductory statistics. Student attitude in two samples was shown to account for four to five per cent of student variability in course achievement. The authors then correlated the subjects' attitude scores with their scores on the Kuder Preference Record and with such variables as: major subject, sex, amount of high school mathematics, amount of college mathematics, number of previous psychology courses and college class. Five variables were selected as the best predictors of student attitudes: The Kuder Computational Scale, the Kuder Persuasive Scale, amount of high school mathematics, amount of college mathematics and number of psychology courses. (The latter negatively related to interest in statistics.) The interpretation is that the greater the number of mathematics courses taken, the greater the familiarity with mathematics, the less the fear and negative attitude toward a statistics course. The negative correlation with number of previous psychology courses is explained as resulting from the fact that this statistics course was a requisite for psychology majors but an elective for students in other fields.

A dissertation by Isard (130) discusses various problems in the construction of attitude inventories toward school subjects. Her conclusions are that: (1) sentence-completion is a superior method to essay-writing for collecting items. (2) Item preference indices are not significantly affected by different instructions to subjects. (3) Ambiguous items are more valid than either socially-acceptable or socially-unacceptable ones. In the traditional questionnaire-type format Isard found a greater tendency for non-achievers to respond in the affirmative (agree) than for achievers to do so, regardless of the direction of preference of the statement. (4) Therefore, this author recommends a forced-choice format on which the subject is asked to pick the *one* statement in a group of four with which he *most* agrees and the *one* with which he *least* agrees. If subjects are asked for the *two* statements with which they *most* agree, again out of a group of four, there is no differentiation between achievers and non-achievers. (5) This form of attitude inventory is impervious to intentional bias.

4. Factors Underlying Attitudes Toward Mathematics

A number of the studies discussed above made some attempt to discover the reasons underlying student preferences. In view of the fact that no real effort was made to assess or evaluate these reasons, they will be mentioned in passing and attention then directed to the few studies which concentrate specifically on determining the factors underlying attitudes.

In the studies on existing preferences toward school subjects, Chase (50) found that when reading was disliked, the reason given was that it was easy, presumably too easy. On the other hand, if arithmetic was disliked, it was considered hard, again presumably too hard. There is a general tendency to find that girls dislike arithmetic more than boys do (50) but Mosher (176) found that sex differences in subject preference did not crystallize until the secondary school level. In the Young study (262), reasons given for losing interest in a subject were: (1) failure to see the need for the course (applied to geometry, algebra, and mathematics as well as to Latin and history); (2) uninteresting; (3) lack of foundation; (4) too difficult. Dutton (75) found that junior high school pupils gave as their reasons for disliking arithmetic that it was too difficult; they did not understand it; it was boring. Those who liked it stressed its practicality and usefulness. He also found that the highest percentage of these pupils named the 5th through 7th grades as being the time at which their attitude toward arithmetic developed. In measuring attitudes of prospective teachers (76), he found a high degree of emotionality accompanying unfavorable attitudes. These subjects stressed their lack of understanding of the subject, boredom, poor teaching and fear of mistakes.

There are various explanations for the factors underlying attitudes toward mathematics or the factors influencing under-and over-achievement. One approach emphasizes the influence of success and failure

PRECEDING PAGE BLANK-NOT FILMED

experiences; another lays the primary influence in the home and correlates student attitude with parental attitude toward mathematics. A third results from a discovery of differences in performance on the basis of sex and attributes these performance differences to different attitudes held by the two sexes which may possibly be related to a masculinity-femininity role image or to other sex differences in nonintellectual variables. This is discussed in more detail in another section of the report (*cf.* Taylor).

Sawin (210) expounds the success and failure approach, claiming that success in mathematics will lead to more positive motivation and failure to negative motivation. From this general hypothesis, he deduces two sub-hypotheses: (1) that there will be a positive correlation between motivation toward mathematics and general mental ability, algebra, geometry aptitude, achievement in terms of mathematics processes and understanding; and (2) that there will be a negative correlation between mathematics motivation and achievement in subjects not related to mathematics.

Sawin's test of motivation is more indirect than any of the other instruments described in the literature. He included three types of items: (1) A list of problems, including mathematics problems for which subjects were asked how often their attention wandered when they had to solve this type of problem. (2) A list of topics (e.g. "How mathematics is used in air navigation") for which subjects were asked if they would like to know more about each topic. (3) Imaginary situations presented ("A classmate wants to tell you about a way of solving a mathematics problem") and subjects asked whether they would be interested.

Sawin then correlated scores on this Motivation Inventory with many other measures. The highest positive correlation was with the V. J. Glennon Test of Basic Mathematics Understandings. However, contrary to his own predictions, he obtained higher positive correlations between the Motivation Inventory and grades in home economics, social studies, and English than between the Inventory and grades in general mathematics. In fact, with mental ability held constant, he obtained a negative correlation between motivation and grades in general mathematics.

Two studies by Gebhard (98,99) examine the effect of success and failure upon the attractiveness of activities. It was found that the attractiveness of a task was determined by past success and failure experiences and also by expectations of future success and failure. Changes in attractiveness were greater, in a positive direction, when the experience of success followed the expectation of failure than when experience and expectation were both of success.

Poffenberger and Norton (195) attribute attitudes toward mathematics to the attitudes expressed by parents. These authors gave an attitude questionnaire to 390 college freshmen, almost equally divided between the two sexes, presumably asking for their attitudes toward school in general. Findings on the attitude questionnaire were that there is a tendency to dislike mathematics. Fifty-two per cent of the students reported liking school in general "very much," but only 25% reported liking arithmetic and mathematics "very much." Two groups of 68 and 75 students were selected, the former liking mathematics "very much" and "better than other subjects," the latter disliking it "some" or "very much." Sex differences came to light in the two groups. The positive group was equally divided between men and women but the negative group consisted of twice as many females as males.

The main purpose of this study was to relate the attitudes of these two groups to various other factors. The only factors which appeared to be significantly related were the attitudes of the parents as reported by the students. Specifically, it was the parents' attitudes toward mathematics and not their attitudes toward school in general which differentiated between the two groups and it was the parents' expectations as to the students' performance in mathematics. Those students who reported that their parents liked mathematics and expected them to do well in it had positive attitudes. The students with negative attitudes on the other hand, said that their parents did not like mathematics and did not expect them to do as well in mathematics as in their other school subjects.

Attitude toward the teacher or influence of the teacher appeared to be operative but not to the same extent as parental influence. In fact, many of the negative group reported liking the teacher but not liking the subject. The negative group consistently rated their mathematics teachers from grade school through high school as poorer than did the positive group and were more critical of them. The authors attribute this to a perseverating negative attitude established in the home.

The final explanation of attitudes towards mathematics or, really, towards problem-solving, results from Sweeney's (232) discovery that there are sex differences in problem-solving ability unrelated to general mental ability, special abilities, or specific knowledge. These sex differences make their appearance in problems requiring a general reasoning factor and restructuring of the material presented.

As a result of this finding, Carey (44) attempted to show that the differential performance of the two sexes was due to a difference in attitude toward problem-solving. She constructed an attitude scale toward problem-solving on which men received significantly higher scores than women. Attitude scores were also found to have some positive relationship to performance scores. A group discussion was held, after a first problem-solving session, to discuss the factors involved in problem-solving success, although the topic of sex differences was assiduously avoided. Following the discussion, the attitude scale and a new set of problems were administered. Female subjects showed a marked improvement in performance following the discussion, although they did not register an improved attitude on the questionnaire. The author concludes that poor problem-solving behavior on the part of women is due to a negative attitude.

This approach is carried one step further by Milton (170) who attributes sex differences in problem-solving to sex-role identification. He administered the Terman-Miles *M-F* Test to 29 college subjects, the *M-F* scale of the MMPI and a test of masculinity-femininity devised by the author called the Behavioral Inventory *M-F*. The subjects were also asked to solve 20 problems on which sex differences had previously been found. Results indicated a positive relationship between masculine sex-role identification and problem-solving skill both across sexes and within a sex, when masculinity-femininity was measured by the Terman-Miles *M-F* Test.

Broadening the approach to a consideration of the non-intellectual personality variables responsible for differences in problem-solving ability, Nakamura (178) attempted to relate conformity to problem-solving skill. One hundred forty-one college subjects were given 20 problem-solving tasks and also a conformity test of the Asch type, that is, in which the subject gives his own judgment of an objective item last, after being told the judgments of four other persons. The extent to which the subject's judgment is affected by the judgments of others is a measure of his conformity. Nakamura found a negative correlation between tendency to conform and achievement in problem-solving when the influence of intelligence was statistically removed.

Berry (17) attempted to re-establish the findings concerning sex differences in problem-solving ability and to establish or re-establish the relationship of this difference to certain non-intellectual variables. These were: attitude toward analytic thinking, conformity, field dependency, anxiety, and masculinity-femininity. Tests administered to the 103 subjects were: the 20 problems used by Nakamura and others, a verbal aptitude test, a mathematics aptitude test, the Terman-Miles *M-F* Scale, the Franck *M-F* Scale, Witkin's Rod and Frame Test, Thurstone and Jeffrey Concealed Figures Test, Carey Attitude Scale, Block Attitude Scale, Sarason Test Anxiety Scale, and a conformity test similar to Nakamura's in procedure but using Asch's lines as the material to be judged.

Results showed that there were differences between men and women on every one of the non-intellectual variables. Men were less anxious, less field dependent, more favorable towards analytic tasks and more masculine. When it came to correlating these differences with differences in problem-solving performance, however, only attitude towards analytic thinking was clearly correlated. Field dependence was related but no correlation found between the Rod and Frame Test and problem-solving for women. Expected correlations were not found for anxiety nor for sex-role identification, except that the Terman-Miles *M-F* was significantly correlated with problem-solving for women.

When the effects of aptitude were removed, only three correlations were significantly related to problem-solving. These three were Concealed Figures, Rod and Frame, Test I, of the field dependency series and the Carey Attitude Scale. Berry's findings thus do not entirely confirm those of either Nakamura or Milton and seem to place greater emphasis upon the relationship of Witkin's field dependency than upon sex-role identification or conformity.

Various unanswered questions remain from these studies.

a. Is there a prevailing negative attitude toward arithmetic and mathematics?

Those researchers who have measured attitudes in small samples, usually of college students (e.g. 76,77,195,262), assert that their subjects held predominantly negative attitudes toward mathematics. The researchers who have attempted to obtain subject preferences of grade-school and high-school pupils from two or three representative schools suggest that mathematics and arithmetic are among the most preferred subjects and that arithmetic is the subject most highly preferred by teachers (50). This latter finding is in contrast to the highly emotional negative attitude data obtained by Dutton from prospective teachers (76).

Taking these data at face value, it is conceivable that arithmetic is not predominantly disliked in grade school but that as the content of mathematics courses increases in difficulty, more and more students develop a negative attitude. From the results obtained by Young (262), mathematics, geometry, and algebra are the courses which, when disliked in high school, have the highest percentage of students who never take a course in this area again. This implies the operation of a strong selective factor such that by the time students reach college, only those students with strong positive attitudes will still be taking mathematics. The remainder have negative attitudes which may very possibly increase in strength with the operation of time and the lack of counteracting influences.

On the other hand, it should be pointed out in criticism of the studies of subject preference that they are inadequate in methodology, usually employing two methods which yield contradictory results (e.g. 176). If it is really of interest to know what the prevailing attitudes are among students, better studies will have to be done, with more adequate questionnaire and sampling techniques.

b. What is the relation between motivation and achievement?

Studies on over-achievers and under-achievers (6,144,200) claim that when mental ability is held constant, differences in achievement are the result of differences in motivation. These studies suffer from an inherent defect, however, in that random errors of measurement on both intelligence and achievement tests, as well as the phenomenon of regression towards the mean, will tend to produce groups of seeming over-achievers and seeming under-achievers. The incidence of true under-achievers and true over-achievers is undoubtedly lower than their apparent incidence resulting from the single administration of an I.Q. and an achievement test to the same subjects.

If one attributes motivation to previous success and failure experiences, as does Sawin (210) and as would many leading theorists, motivation and ability cannot be regarded as entirely independent variables. Since ability is the best predictor of achievement, and therefore of success, positive motivation resulting from previous rewarded experience must be correlated with ability. This explains the discrepancy between the results obtained by Sawin, who found that when mental ability was held constant the correlation between motivation and achievement was negative, at least on one measure, and the results obtained by the other investigators in this area. A certain portion of motivation must be correlated with ability but, over and above this, there are differences in motivation, unrelated to ability, which discriminate among over- and under-achievers of equal intellectual potential.

Another difficulty with this type of study is the unreliability of the instruments used. The motivational inventories may have such a high degree of unreliability that when correlating them with grades obtained (which are also unreliable), the extent of correlation is necessarily reduced. When, then, the researcher controls for mental ability, he is reducing an already unreliably low correlation by the necessarily much higher correlation between mental ability and grades achieved.

c. What are the factors underlying attitudes toward mathematics?

The same problem which was raised in the previous discussion of the relation between attitude and achievement is relevant again. Is motivation the result of previous success and failure experiences, or is it governed by parental attitudes or such personality traits as conformity, sex-role identification, field-dependence, etc.?

Learning theorists would predominantly agree that positive motivation results from rewarded or reinforced experience. And such experiences, in the case of success or failure in mathematics, would be highly dependent upon ability. However, over and above this, there are undoubtedly other motives which could receive reinforcement in mathematics training, namely, desire for parental approval, peer approval, fitting into the sex-role prescribed by one's society. Consequently, in order to answer the question of factors underlying motivation, it would be necessary to disentangle that portion correlated with ability from that portion which is uncorrelated but which might be correlated with other factors. This has been attempted in the studies on sex-differences in problem-solving but without conclusive results as yet.

C. Personality Factors

A number of references discuss the relation between personality and learning. Signori (219) claims a relation between personality and thinking processes; Slinger (221) investigated personality problems of college freshman and found an inverse relation between adjustment and achievement. Discussions of hostility points out that classroom competition promotes anxieties and repressed hostility (64) and that different classroom levels of permissiveness toward the expression of hostility affect learning and performance (103).

Discussion of under-- and over-achievers in terms of personality traits, rather than motivation, are frequent in the literature. Cooper (36) found under-achievers less self-analytic, more careless, less self-critical and with lower grade-point aspirations than achievers. Bloomberg (25) characterizes the typical college achievers as having a definite goal, believing grades to be important, being superficially introverted, secure, self-confident and with mild anxieties. A not too surprising finding of Berger and Sutker (15) is that higher intelligence and personality adjustment result in higher academic performance.

An attempt by Tibbetts (242) to relate family relationship to achievement found a striking degree of disagreement between under-achievers and parents as to standards of behavior, in contrast to a higher incidence of agreement for high achievers. The variable of self-concept has come under investigation: Reeder (208) found that children with low self-concepts achieve lower in relation to their potential. High-achievers are found by Rabinowitz (198) to suffer from doubt and confusion in the area of family and peer acceptance, a factor which may push them to strive harder in the hopes of academic acceptance. A study of under- and over-achievers by Shimmon (217) found high achievers to be superior in high school marks, study habits, analytic thinking, industriousness and attitude toward mathematics. They were more highly regarded by teachers and more interested in occupations of an abstract or intellectual nature. Almost all planned professional careers and a significant number had parents who were foreign-born. The under-achieving group had a more favorable attitude toward social studies, were more interested in manual skills and working with people.

Some studies specific to achievement in reading suggest that over-achievers had significantly better mental health, significant differences between under- and over-achievers lying in the areas of immaturity, emotional instability, feelings of inadequacy and attitudes about physical defects (22). Over-achievers were primarily introversive and under-achievers extratensive with a great capacity to respond to environmental stimuli. Abrams (1) suggests that non-readers experience difficulty in attention as a result of anxiety. They are more impulsive and less able to respond appropriately to emotional stimuli. A contrary hypothesis concerning anxiety is suggested by Lynn (153) who compared anxiety levels among school boys and their discrepancy scores between achievement in reading and achievement in arithmetic. He found a tendency for anxiety to be associated with good reading and suggests that perhaps the highly anxious child spends more time in solitary reading whereas his high anxiety interferes with the performance of a complex task such as arithmetic.

On the personality characteristics related to mathematics achievement, Egner & Obelsky (80) suggest that students who score high on a sterotypy test are better at mathematics and science, whereas low scorers are superior in the humanities and social sciences. These findings are of dubious validity in view of the lack of statistical treatment of results and the questionable character of the items included on the stereotypy test. An investigation of the relative contribution to achievement in college mathematics of work habits as against aptitude was carried out by Krathwohl (141). For the group as a whole, prediction of achievement was found to be independent of work habits and the contribution of industriousness was negligible. When the group was divided into above average, average and below average groupings on the variable of industriousness, the most accurate prediction of achievement could be made on the basis of aptitude.

An unpublished dissertation by Rochlin (207) represents a detailed investigation of the personality variables associated with achievement in a college freshman course in general mathematics. The significant characteristics of under-achievers in the mathematics course were found to be: an absence of both liking for and confidence in handling mathematics; an approach to non-interpersonal activity (such as tests, puzzles, etc.) characterized by non-flexible organization, lack of self-confidence, non-positive cathection; conflict or emotional disturbance associated with awareness of inadequacy; the predominant use of an intuitive rather than a deductive approach; lack of self-confidence in relations with instructors, and the prevalent prediction of unfavorable outcomes for self and peers in school situations. Over-achievers, on the other hand, were characterized by: an absence of dislike for mathematics; flexible organization and persistence in non-interpersonal activity; persistence in the face of obstacles to school accomplishment and to personal growth for greater social and emotional maturity; the inclusion of abstract considerations, and the prediction of favorable outcomes for the self in school situations.

A more clinical or depth-psychology approach to the study of the relation of personality traits to mathematics behavior may also be found. One study by Plank (192), involving case work with five children ages 8-9 who were retarded in mathematics, attributes their difficulties to anxiety, insecurity, family position (either only children or much younger or much older than other siblings), maternal overprotection, and rigidity. Another study (193) of the emotional components of arithmetic learning as seen through autobiographies suggests that, for males, the over-protected child is poor in arithmetic while relatively very good at reading. Famous mathematicians, and non-mathematicians with ability and interest in mathematics, on the other hand, were found to have had traumatic childhood experiences with their parents. The authors feel that aggression is also necessary to mathematical activity; aggression is harnessed and turned against abstractions rather than people. Women's interest in mathematics is tied to identification with a strong male figure (who is himself interested in mathematics) and with lack of feminine identification or conflict in this area.

Two somewhat different psychoanalytic approaches to mathematics were found. One suggested that the unconscious is capable of mathematical operations distinct from the conventional kind (8); the other offered an analysis of the fantasy productions of a 21-year-old, disturbed mathematician (208).

D. Intelligence and Special Abilities

The area of mental abilities is of direct relevance to mathematics learning. The question of whether proficiency in mathematics is due to a general intelligence factor or to special aptitudes, or both; of whether the study of mathematics contains topics which depend upon the same underlying factor, that is, whether it is a unitary subject matter or not, all have come under experimental investigation. Many of these studies were performed prior to the 10-year period covered by this survey and only a few of these earlier investigations are mentioned here.

A developmental approach to the study of mental abilities is found in an investigation by Doppelt (69). The author hypothesized a reduction in the amount of undifferentiated general ability, an increase in the specialization of ability components, with increasing age. He found a general factor closely associated with reasoning ability, a factor which did not decrease in importance among the subjects he studied, who were 13- to 17-year-olds. This is an elaborate factorial study performed on a large population and with a review of other findings in the same area.

The nature of mathematical ability has come under considerable scrutiny. A study by Blackwell (23) found a central intellectual factor in mathematical ability associated with the power of selective, quantitative thinking and of deductive reasoning involving the ability to apply general principles to particular cases, to abstract and to generalize and to make deductions. A second factor is described as an "operation in imagery" factor involving visual and spatial imagery. A third factor is verbal and a fourth, found only in girls, involved exactness and precision. An early investigation by Fleck (88) of mathematical abilities of school boys found the ability of an individual in one branch of the subject to be a good indication of his ability in other branches. An analysis of certain components of mathematical ability by Coleman (54) found computational ability to be distinct from and only remotely related to mathematical ability, whereas the latter was closely akin to ability in deductive reasoning. A good memory was not found to be essential to mathematical ability, and the ability to see spatial relationships was only helpful in a limited area.

Rusch (209) performed a factor analysis of arithmetic achievement in grades 4, 6 and 8 and found that the number factor has three sub-factors: abstraction, analysis, and application. Boys did not seem to develop analysis or abstraction factors as clearly as girls and the 6th grade appeared to be the key grade in which unusual changes in factorial make-up occurred. McAllister (155) claims to have found a general factor underlying mathematical ability, as well as special factors. The general factor is an aspect of general intelligence: the capacity to operate with arithmetic concepts. Similar results were reported by Barakat (9) with the finding that basic or general ability or innate intelligence, plays the largest role in all mathematics attainment. When this basic factor was removed, special factors made their appearance in the following order of importance: a) mathematical reasoning; b) symbolic factor.

In a study comparing retarded and normal pupils in mathematics, Hamza (113) concluded that mathematics is an integrated subject since retardation occurred in all three branches studied (arithmetic, algebra and geometry). The most important factor in mathematical ability was found to be the general intellectual factor. A second factor of some importance was that of visual imagery, in which the retarded pupils were deficient. A study by Pitts (191) of the relationship between functional competence in mathematics, reading grade levels, intelligence, and age found a significant correlation between mathematical and reading abilities. A study of the changes occurring in three factors over time (235) resulted in the finding that between the 6th and 9th grades the three factors tended to remain the same but the general factor showed the greatest increase over time, the verbal factor increased by a smaller amount, and there was no apparent increase in the spatial factor.

Interest in the possibility of altering performance on achievement or intelligence tests is found in two studies. Tilton (243) told teachers to attempt to raise the arithmetic scores of their students on an achievement profile, and found that there was a significant improvement for 10 of the 37 experimental subjects, although not a significant improvement for the experimental group as a whole. Watts (255) explored the finding that mathematics students tend to have higher scores on intelligence tests than do other subjects. She found that while, on the first testing, university mathematics students did have higher scores than non-mathematicians, the improvement in score at a second testing one year later was only slightly greater for mathematics students. A similar investigation of practice effects in the repeated administration of intelligence tests is reported by Peel (187).

Studies exist regarding specific mathematical abilities. Swineford (234) examined the nature of a number factor. Murray (177) investigated the contributions of certain primary abilities to achievement in geometry. The primary abilities in question were spatial relations, reasoning, numerical and verbal abilities. Spatial ability was found to have the least importance in accounting for the variance in geometry performance, as measured either by class grades or a geometry achievement test. Two studies of abilities underlying performance in arithmetic problem-solving found that the relationship between general intelligence and selected problem-solving abilities was much higher than the relationship between achievement in mathematics and these abilities (152), and also that vocabulary and reasoning abilities appear to be the most important abilities contributing to the solving of arithmetic problems (132).

An investigation of the relation of reading abilities to success in mathematics (79) suggests that while general vocabulary aptitude is highly correlated with mental age it has a low correlation with success in mathematics. Three groups of subjects were found: 1) those low in mental age and low in reading comprehension, among whom the slow readers were also poor in mathematics; 2) those of average mental age and average reading comprehension, with slow readers excelling in mathematics; and 3) those of high mental age and high reading comprehension, with fast readers excelling in mathematics. A suggested explanation is that the factor associated with slow reading in the first group also contributes to poor mathematics performance. The slow readers in the second group excelled in mathematics perhaps through the habit of careful work, whereas the fast readers in the third group were superior performers in all respects.

A final study examines the relationship between concept formation and reading achievement (142). Among subjects matched for intelligence, non-readers, as compared to readers, were found to lack versatility, flexibility and originality in devising suitable hypotheses for testing, and also to lack initiative and persistence in verifying solutions. They were also deficient in ability to draw inferences from relevant clues, ability to shift set, to analyze factors, to devise adequate labels and to form adequate concepts for dealing with language. The non-readers also exhibited a tendency to cling to previous solutions, an abnormal need for success and avoidance of failure, a dependence upon the physical characteristics of objects and a tendency to be more concrete and less abstract in conceptual functioning.

In summary, studies in this area appear to agree on the importance of a general intellectual factor for ability in mathematics but the investigation of specific abilities is not conclusive nor, perhaps, is the approach to this problem as meaningful as it could be. Two research designs presented in another section of this report suggest somewhat different approaches. Gagné proposes to examine the effect upon mathematics learning of different instructional procedures designed to fit in with certain special abilities. Feierabend raises the question of whether differences in type of special abilities might facilitate or inhibit mathematics learning for persons of equal general intelligence. The unanswered question remains as to whether all persons of sufficient general intelligence have equal potential for mathematics, or whether there may not exist some special abilities, factors or conceptual approaches which are specific to the field of mathematics or perhaps to creativity in mathematics thinking. This leads naturally into the next topic of study, which is concerned with the thought processes underlying mathematics behavior.

E. Thought Processes

One area of interest in the investigation of thought processes is the developmental approach to child thinking. Much of the work in this area is by Piaget. A recent article discussing the formation of geometrical concepts in children (189) suggests that the earliest form of geometric understanding is topology. At three, the child distinguishes between open and closed figures. Only later does he approach an understanding of projective (perspective) and Euclidean geometry. Around six or seven years the child grasps the principle of conservation of quantity, which is a prerequisite to the concept of number. Discovery of reversibility is a prerequisite to an understanding of geometrical concepts. Measurement is viewed as a derived concept which, when understood, leads to the concepts of division and substitution. Measurement in two or three dimensions is understood at about eight or nine years of age. This leads to a comprehension of the fundamentals of Euclidean geometry. Horizontality and verticality are not understood until the age of nine. When the child has discovered how to construct coordinate axes by reference to natural objects, which he does at about the same time as he conceives of coordination of perspectives, he has completed the conception of how to represent space. The child has now developed the fundamental mathematical concepts which spring spontaneously from his own logical operations. Two articles by other authors (151,163) stress the necessity for children to approach the study of arithmetic so as to grasp the abstract nature of number concepts. "Three-ness" must be understood divorced from its concrete instances. Two studies on developmental trends in abstraction ability (218) and arithmetic ability (128) suggest that topics in arithmetic are placed in the curriculum without regard for the type of material that the children are able to understand at a given age.

A group of studies has attempted to analyze the thought process involved in various aspects of arithmetic. Gibb (102) examined children's thinking in the course of subtraction. In a factorial design comparing three problem applications of the subtraction process (take-away, additive and comparative) with three contexts (concrete, semi-concrete and abstract), she found that relatively young children with limited number experience give different responses to problems involving the three applications of subtraction. They do not understand that one basic idea underlies all applications but respond to each problem in terms of the reality of the situation. Also, the contexts used suggest different responses. A similar study of thought processes has been made for the learning of multiplication and division problems (109) and for higher decade addition (189). The method employed in the latter study was to ask for the children's self-reports on their thought processes or else to suggest methods and have them select the one which best corresponded to their own thinking.

The thought processes involved in scientific analysis and relational thinking have also been studied. Chess (51) examined the manner of analysis employed in three types of experimental problems and found inter-problem fluctuation characteristic of all subjects as well as common analytic features and inter-individual differences in thought. A dissertation by Jackson (131) attempted to develop the ability to interpret data in ninth grade subjects by the use and understanding of algebraic concepts. He found a significant improvement especially in recognizing the limitations of data following his instruction, greater improvement being made by students of higher intelligence. A study by Burch (38) of the value of instruction in formal analysis for proficiency in problem-solving found that pupils tended to score higher on the test which did not require them to go through the steps of formal analysis, even though they had been trained in formal analysis. They did, however, consistently answer the "what must you do?" step correctly prior to giving a correct solution.

A few analyses have been made of the types of processes involved in concrete and abstract thinking(119) and objective thinking abilities (165). A most interesting work is the discussion of invention in the mathematical field by Hadamard (110). As a mathematician himself, he analyzes his own experiences and those of other mathematicians on an intuitive and introspective level. Emphasis is laid upon the functioning of unconscious processes in solving especially baffling problems, and also upon the existence of different types of mathematical minds and of different modes of conceptualization.

An attempt has been made by McNemar (158) to differentiate between college students of high and low reasoning ability on the basis of their thought processes. The three variables studied were: 1) free and controlled word association; 2) deduction and induction; and 3) experimentally induced set. Two main findings were that the high reasoning ability group was superior in speed of controlled association and in ability to overcome experimentally induced set.

An investigation of the retention of plane trigonometry by Celauro (49) related the forgetting of this subject by college freshman completing a course to the time elapsed before testing, the intelligence of the subjects, their mathematical proficiency and the amount initially learned. Findings were that the upper one-third in achievement on the final examination retained more than the lower one-third. However, in terms of per cent retained of amount originally learned, the difference between upper and lower thirds was insignificant. Thus, generalizations concerning retention depend upon the definition of retention employed. With either definition, however, the author concludes that retention is not sufficient for students to do satisfactory work in other mathematics courses for which trigonometry is a pre-requisite.

The question of the transfer of learning to new situations is of major interest to mathematics. There are two aspects to this question. One concerns the often-made statement that one of the values of an understanding of mathematics is that it transfers to the study of other subjects; it teaches logic and reasoning ability. No studies were encountered specifically directed to this problem. A review of the role of transfer in the teaching of mathematics may be found in several articles on the relation of learning theory to the teaching of mathematics (68,133,149).

A study by Brownell (34) examined the effects of practicing a complex skill (division by two-place numbers) upon the learning of constituent sub-skills. He found both positive and negative results. There seemed to be evidence of an inhibitory as well as of a facilitating effect, depending upon the degree of similarity between the two tasks and the extent of prior learning of the sub-skill.

Three studies compared the relative efficiency for transfer to new tasks of learning by rote versus learning by understanding. Carpenter (47) compared the relative effectiveness of rote and functional learning on a post-test involving the identification of new material. Functional learning was found superior for this purpose, although rote learning was more efficient initially. Studies by Hilgard et al. (125) examined the relative efficiency of learning by memory or learning by understanding for transfer to new tasks. Using Katona's card-trick materials, these authors found in their first study that more time was required for initial learning by the understanding group, and that overnight retention was equal for both the understanding and memorization groups. The understanding group did show a slight advantage in transfer but continued to make a large number of errors. In an attempt to cut down on these errors, the second study was performed. Subjects in the understanding group were taught by five different methods but the over-all differences in success among the methods were slight. This last study revealed significant sex differences in the ability to transfer, with superior performance by the boys. This finding is supported by Kostick (139).

Another approach to the problem of transfer is the question of whether, in learning a principle, it is better to have it in verbal or non-verbal form and whether it is better to impart a knowledge of the principle to the subjects or to let them achieve their own understanding. A dissertation by Corman (58) compared three degrees of information (none, some, and much) applied to two aspects of problem-solving, principle of solution and method of solution. The three measures of effect were number of tasks solved, transfer to simple and complex problems, and verbalizing of the principle. Results showed that information concerning the principle was not effective for problem-solution, whereas success in problem-solving increased with degree of information about method for subjects of high intelligence. Success in verbalizing the principle, however, was uncorrelated with success in problem-solving, and increased with amount of information as to principle.

A study by Hendrix (121) assessed the relative value of three methods of teaching general principles for transfer to similar problems. Method I involved presenting the principle to the students with illustrations. Method II was an unverbalized awareness procedure: the subjects worked on simple examples until they became aware of the generalization but the generalization was never verbalized. Method III was the same as Method II but the subjects were also asked to verbalize the principle. Results showed Method II to be superior. Method III was next in effectiveness and Method I least effective. In support of this finding, Haslerud and Meyers (116) compared the transfer value on a coding task of subject-derived or experimenter-given principles. The advantage clearly lay with the subject derived principle. In another section of this report, Hendrix has suggested a modification of the Haslerud and Meyers design.

It is difficult to summarize the area of thought processes, since it covers such a wide variety of topics. The question of the developmental understanding or readiness to grasp mathematical concepts has not as yet been put to experimental investigation. Among the research designs contained in another section of this report, one suggests various procedures for a more experimental approach in this area (see Rosenbloom).

The problem of transfer is another area in which more work is needed. While the question of verbalization or non-verbalization of general principles seems to have received experimental attention, the transfer value of mathematics to the study of other subjects or the transfer in learning from one topic to another within the field of mathematics is still virgin territory. Unfortunately, a prior understanding of the nature of transfer, of the inhibitory and facilitating effects of similarity between tasks and, as a further refinement, of similarity between responses versus similarity between stimuli, is still necessary and will depend upon further research. A theoretical model for the breakdown of stimuli and responses on a continuum of similarity-dissimilarity may be found in Shepard (215).

Another question is the evaluation of transfer. Assessing the increase in performance on one task due to prior performance on an earlier task is much the same as assessing learning and involves the same problems discussed in that section above. The point was raised in discussion at the Conference covered by this report that the use of a residual gain score (cf. 71, 162) for the evaluation of transfer might be as applicable as for the evaluation of learning.

IV. EVALUATION AND TESTING

Many of the problems involved in the evaluation of learning have been discussed in various sections above. One question to which considerable attention has been devoted in the literature is the construction of reliable and valid tests of mathematical proficiency and aptitude. An over-all review of tests in the mathematical field may be found in the series of Mental Measurements Yearbooks (39,40).

Frederiksen and Satter (91) attempted to construct a test of arithmetical computation which would make few if any demands on verbal ability. Despite this aim, a substantial correlation was found between test scores and scores in a measure of verbal ability. Various studies examine the usefulness of certain aptitude tests as predictors of proficiency in mathematics. Grime (105) found the degree of correlation between the Iowa

Aptitude Test and success in elementary algebra to be about .70. Kvaraceus and Lanigan (145) found the lowest mean correlation between the Otis Quick Scoring Mental Ability Test and the Iowa Every-Pupil Tests of Basic Skills to be in the area of arithmetic. This correlation was .59 for arithmetic, whereas for reading comprehension it was .78, the highest mean correlation.

Sutherland (229) found a marked drop in performance on a test of arithmetic attainment during the first three years in a British six-year secondary school, in comparison to the performance of the same pupils on the same test at the time of their transfer from primary school. Possible explanations are that more time is spent on arithmetic in the primary schools, or that the pupils are more motivated at the time of taking the transfer examinations, producing scores which actually overestimate their abilities. In another study, Sutherland (231) assessed the prognostic value of certain arithmetic tests and found that the problem test was the best predictor of secondary school success. Northrup et al. (180) advocate the use of functional item analysis to detect need for remedial help in mathematics.

The construction of tests of arithmetic readiness is a topic of interest to various investigators. Brueckner (37) believes that tests of arithmetic readiness should not serve as predictors but as diagnostic devices for the determination of factors likely to interfere with learning. He devised and validated a readiness test for easy work in division by two-figure divisors. Brownell (32) also devised a readiness test for work in division and, after administering it to children in twenty 5th-grade classes in three school systems, concluded that approximately one-fifth of the students were starting a new type of division with serious deficiencies in accuracy of computation in basic skills.

A test of arithmetical concepts and abilities of kindergarten, first and second grade children was devised by Carpenter (46), covering the areas of counting, measurement and computation. In analyzing performance, he found significant sex differences (girls superior in total test score), rural-urban differences, and differences correlated with father's occupation. In terms of validation, there were also significant differences between total test scores and teacher's marks for all grades. Olander, Van Wagenen and Bishop (181) found that ability in arithmetic in first grade, as measured by scales of quantitative information and perception of quantitative relations, can be used to predict performance in arithmetic three years later.

Collier (55) has designed a test to measure objectives now commonly neglected in evaluation, namely, arithmetic understanding and reasoning ability, and Reilly (203) has devised a technique to measure a component of critical thinking in the area of physics.

Tests, either of a standard kind or of the author's devising, have been administered to determine the level of proficiency in mathematics of different segments of the population. Students tested in six colleges by Hannon (114) were found to give better performances on a general mathematics test with increased mathematics training, a not-too-startling result. Miller (169) administered an arithmetic proficiency test to college freshman and found significantly more females with low scores on the arithmetic test than males. Education students received low scores most frequently; engineering students least frequently. Multiple deficiencies (reading, writing and arithmetic) occurred most frequently with deficiency in writing and least frequently with deficiency in arithmetic. Miller also found the lowest college attrition (drop-out) rate among students with high scores in arithmetic proficiency and the highest over-all scholastic achievement among this group.

An investigation of teachers' understanding of arithmetic carried out by Orleans and Wandt (183) used tests devised by the authors to cover a representative sampling of basic arithmetic skills, processes and concepts. The authors feel that there are relatively few processes, concepts or relationships which are understood by a large percentage of teachers.

A somewhat different focus is shown by Kramer (140) who compared the arithmetic achievement of school children in the U.S. and in the Netherlands. While Dutch children appear to be two years ahead of American children in arithmetic, this may be explained by the greater selectivity of students in Holland and the greater length of time spent on the study of arithmetic in school. If the length of time spent on the material is equalized, high ability Dutch and high ability American pupils are equal in performance.

A more comprehensive discussion of the evaluation of learning is contained in another section of this report (see Gleser).

V. RESEARCH IN OTHER COUNTRIES

The final section of this survey is intended as a brief reminder of the fact that much research on the teaching and learning of mathematics is being carried out in countries other than the United States. A recent review of foreign research (264) illustrates this point. Apart from the English-speaking countries, such as Great Britain, Australia and New Zealand, considerable work is being done in France, Belgium, Germany, Sweden, and Japan. A listing of selected foreign references will be found at the end of the bibliography, but this does not duplicate the listings to be found in the AERA review (264).

In very cursory survey, it seems that work reported from the English-speaking countries, from Sweden and from Japan consists primarily of research on specific problems in the teaching of arithmetic, the nature of errors, diagnosis and arithmetic readiness. The work reported from French- and German-speaking countries, on the other hand, shows a greater interest in the thought processes involved in performing arithmetic operations, the development of concepts and of abstraction and analytic thinking, as well as of space experience and functional thinking.

BIBLIOGRAPHY

1. Abrams, J. C. A study of certain personality characteristics of non-readers and achieving readers. *Dissertation Abstr.*, 1956, *16*, 377-378.
2. Adams, Olga. Arithmetic readiness in primary grades. *Suppl. educ. Monogr.*, 1947, No. 63, 10-16.
3. American Educational Research Association. The natural sciences and mathematics. *Rev. Educ. Res.*, 1951, *21*, No. 4.
4. American Educational Research Association. Twenty-five years of educational research. *Rev. Educ. Res.*, 1956, *26*, No. 3.
5. Anderson, G. L. Quantitative thinking as developed under connectionist and field theories of learning. In Swenson, E. J., Anderson, G. L., and Stacey, C. L. *Learning theory in school situations*. Minneapolis: U. of Minnesota Press, 1949.
6. Armstrong, M. E. A comparison of the interests and social adjustment of under-achievers and normal achievers at the secondary school level. *Dissertation Abstr.*, 1955, *15*, 1349.
7. Atkinson, J. W., and Raphelson, A. C. Individual differences in motivation and behavior in particular situations. *J. Personality*, 1956, *24*, 349-363.
8. Baker, S. J. The mathematics of the unconscious. *J. clin. exper. Psychopath.*, 1951, *12*, 192-212.
9. Barakat, M. K. Factors underlying the mathematical abilities of grammar school pupils. *Brit. J. educ. Psychol.*, 1951, *21*, 239-240.
10. Basescu, S. Learning and intelligence: a study of the relationship between intelligence and reinforcement. *Dissertation Abstr.*, 1954, *14*, 181.
11. Bebb, A. M. A study of two methods of teaching higher decade addition. *Dissertation Abstr.*, 1953, *13*, 730.
12. Bebb, R. R. Comparison of a tachistoscopic presentation with a classroom method of teaching the basic multiplication facts. *Dissertation Abstr.*, 1952, *12*, 693-694.
13. Bendig, A. W., and Hughes, J. B. Student attitude and achievement in a course in introductory statistics. *J. educ. Psychol.*, 1954, *45*, 268-276.
14. Berdie, R. F. Aptitude, achievement, interest, and personality tests: a longitudinal comparison. *J. appl. Psychol.*, 1955, *39*, 103-114.
15. Berger, I. L., and Sutker, A. R. The relationship of emotional adjustment and intellectual capacity to academic achievement of college students. *Mental Hygiene*, 1956, *40*, 65-77.
16. Bernstein, A. L. A study of remedial arithmetic conducted with ninth grade students. *Dissertation Abstr.*, 1955, *15*, 1567-1568.
17. Berry, P. C. *An exploration of the interrelations among some non-intellectual predictors of achievement in problem-solving*. Technical Report 4, Office of Naval Research, 1958.
18. Betz, W. Functional competence in mathematics - its meaning and attainment. *Math. Teach.*, 1948, *41*, 195-205.
19. Betz, W. The transfer of training, with particular reference to geometry. *Nat. Council of Teachers of Math. Yrbk.*, 1930, *5*, 149-198.

20. Betz, W. Teaching and learning processes in mathematics. *Math. Teach.*, 1949, *42*, 49-55.

21. Billig, A. L. Student attitude as a factor in the mastery of commercial arithmetic. *Math. Teach.*, 1944, *37*, 170-172.

22. Blackham, G. J. A clinical study of the personality structures and adjustments of pupils under-achieving and over-achieving in reading. *Dissertation Abstr.*, 1955, *15*, 1199-1200.

23. Blackwell, A. M. A comparative investigation into the factors involved in mathematical ability of boys and girls. Part I and Part II. *Brit J. educ. Psychol.*, 1940, *10*, 143-153; 212-222.

24. Bloom, B. S. Implications of problem-solving difficulties for instruction and remediation. *Sch. Rev.*, 1947, *55*, 45-49.

25. Bloomberg, M. The prediction of scholastic success through the use of a forced-choice problems- and-attitudes inventory. *Dissertation Abstr.*, 1955, *15*, 2566.

26. Bond, J. G. A study of the comparative effects of positive and negative stimuli in a non-verbal learning situation. *Dissertation Abstr.*, 1955, *15*, 226-227.

27. Bouchard, J.B. An exploratory investigation of the effect of certain selected factors upon performance of sixth grade children in arithmetic. *J. exp. Educ.*, 1951, *20*, 105-112.

28. Brown, F. R. The effect of an experimental course in geometry on ability to visualize in three dimensions. *Dissertation Abstr.*, 1955, *15*, 83-84.

29. Brown, K. E. *Analysis of research in the teaching of mathematics 1955 and 1956*. Bull. 1958, No. 4. Washington, D.C.: U.S. Dept. of Health, Education and Welfare, Office of Education, 1958.

30. Brown, K. E. *Mathematics Education research studies-1953*. Circular No. 377-11. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1954.

31. Brown, K. E. Research in teaching high school mathematics. *Math. Teach.*, 1958, *51*, 593-595.

32. Brownell, W. A. Arithmetic readiness as a practical classroom concept. *Elementary Sch. J.*, 1951, *52*, 15-22.

33. Brownell, W. A., and Moser, H. E. *Meaningful vs. mechanical learning: a study in grade III subtraction*. Durham, N. C.: Duke University Press, 1949.

34. Brownell, W. A. The effects of practicing a complex arithmetical skill upon proficiency in its constituent skills. *J. educ. Psychol.*, 1953, *44*, 65-68.

35. Brownell, W. A. Frontiers in educational research: in arithmetic. *J. educ. Res.*, 1947, *40*, 373-380.

36. Brueckner, L. J. The development of readiness tests in arithmetic. *J. educ. Res.*, 1940, *34*, 14-20.

37. Brueckner, L. J. The development and validation of an arithmetic readiness test. *J. educ. Res.*, 1947, *40*, 496-502.

38. Burch, R. L. Formal analysis as a problem-solving procedure. *J. Educ.*, 1953, *136*, 44-47.

39. Buros, O. K. (ed.) *The third mental measurements yearbook*. New Brunswick, N. J.: Rutgers U. Press, 1949.

40. Buros, O. K. (ed.) *The fourth mental measurements yearbook*. Highland Park, N. J.: The Gryphon Press, 1953.

41. Buswell, Margaret M. The relationship between the social structure of the classroom and the academic success of the pupils. J. exp. Educ., 1953, 22, 37-52.
42. Butler, C. H., and Wren, F. The teaching of secondary mathematics. New York: McGraw-Hill, 1941, 1st ed.
43. Capobianco, R. J. Studies of reading and arithmetic in mentally retarded boys. II. Quantitative and qualitative analyses of endogenous and exogenous boys on arithmetic achievement. Monog. Soc. Res. Child. Development, 1954, 19, (1), 101-142.
44. Carey, Gloria L. Sex differences in problem-solving performance as a function of attitude differences. J. abnorm and soc. Psychol., 1958, 56, No. 2.
45. Carlson, C. R. A study of the relative effectiveness of lecture and and directed discussion methods of teaching tests and measurements to prospective Air Force instructors. Dissertation Abstr., 1953, 13, 1112.
46. Carpenter, C. L. Identification and measurement of the arithmetical concepts and abilities of kindergarten, first and second grade children. Dissertation Abstr., 1957, 17, 2205-2206.
47. Carpenter, F. The effect of different learning methods on concept formation. Science Educ., 1956, 40, 283-285.
48. Caton, W. B. The teaching of mathematics. Sch. Sci. and Math., 1954, 54, 490-492.
49. Celauro, E. L. Factors associated with the retention of plane trigonometry. Dissertation Abstr., 1953, 13, 100-101.
50. Chase, W. L. Subject preference of fifth grade children. Elem. School J., 1949, 50, 204-211.
51. Chess, Edith G. The manner in which two samples of ninth-grade general science students analyze a number of selected problems. Dissertation Abstr., 1956, 16, 1398-1399.
52. Childs, G. B. Success in initial university mathematics courses of students with correspondence and non-correspondence background in high school mathematics. J. educ. Res., 1956, 49, 607-612.
53. Clark, Eileen. Number experiences of three-year-olds. Childh. Educ., 1950, 26, 247-250.
54. Coleman, R. H. An analysis of certain components of mathematical ability and an attempt to predict mathematical achievement in a specific situation. Dissertation Abstr., 1956, 16, 2062.
55. Collier, C. C. The development and evaluation of a non-computational mathematics test for grades 5 and 6. Dissertation Abstr., 1957, 17, 1027-1028.
56. Cooper, M. N. To determine the nature and significance, if any, of certain differences in the social and personal adjustment of fifty-one non-successful college students at Texas Southern Univ., Dissertation Abstr., 1956, 16, 497.
57. Corle, C. G. The characteristics of the procedures of good and poor problem solvers in sixth grade arithmetic. Univ. Cincinnati, Abstr. Grad. Theses Educ., 1944-1954, 5, 218-227.
58. Corman, B. R. The effects of varying amounts and kinds of information as guidance in problem-solving. Dissertation Abstr., 1954, 14, 1176-1177.
59. Counts, Sarah. Achievement in college mathematics as a function of instructors' and students' patterns of primary mental abilities. U. Chicago, Ph.D., 1952.

60. Cruickshank, W. M. Arithmetic ability of mentally retarded children. Ability to differentiate extraneous materials from needed arithmetic facts. *J. educ. Res.*, 1948, *42*, 161-170.

61. Cruickshank, W. M. Arithmetic work habits of mentally retarded boys. *Amer. J. Mental Def.*, 1948, *52*, 318-330.

62. Curry, J. F. The effect of reading instruction upon achievement in seventh grade arithmetic. *Dissertation Abstr.*, 1955, *15*, 2059.

63. Dattman, P. E., and Israel, H. E. The order of dominance among conceptual capacities: an experimental test of Heidbreder's hypotheses. *J. Psychol.*, 1951, *31*, 147-160.

64. Davidson, H. A. Competition, the cradle of anxiety. *Education.*, 1955, *76*, 162-166.

65. Davis, C. E. The association of subject preference with accomplishment in selected scholastic areas of eighth grade pupils. *Dissertation Abstr.*, 1956, *16*, 1225-1226.

66. Davis, D. J. *A comparative study of achievement levels of twelfth-grade pupils on a test designed to measure functional competence in mathematics*. Unpublished Ph.D. Dissertation, Univ. of Mich., 1950.

67. Deans, Edwina. The effect of certain immature procedures on the learning of arithmetical processes by second-grade children. *Univ. Cinn., Abstr., Grad. Theses. Educ., 1944-1954*, 1955, *5*, 129-147.

68. DeLong, A. R. Learning. *Rev. Educ. Res.*, 1955, *25*, 438-452.

69. Doppelt, J. E. The organization of mental abilities in the age range 13 to 17. *Teach. Coll. Contr. Educ.*, 1950, No. 962.

70. Douvan, Elizabeth. Social status and success strivings. *J. abnorm. soc. Psychol.*, 1956, *52*, 219-223.

71. DuBois, P. H., and Manning, W. H. (eds.) *Methods of research in technical training*. Technical Report No. 3, Office of Naval Research Contract No.Nonr 816(02), Washington University, May, 1957.

72. Duncker, K. On problem-solving. *Psychol. Monogr.*, 1945, *58*, No. 5.

73. Dunn, L. M. Studies of reading and arithmetic in mentally retarded boys. I. A comparison of the reading processes of mentally retarded and normal boys of the same mental age. *Monogr. Soc. Res. Child Developm.*, 1954, *19* (1), 7-99.

74. Durr, W. K. The use of arithmetic workbooks in relation to mental abilities and selected achievement levels. *Dissertation Abstr.*, 1955, *15*, 2126.

75. Dutton, W. H. Attitudes of junior high school pupils towards arithmetic. *Sch. Rev.*, 1956, *64*, 18-22.

76. Dutton, W. H. Attitudes of prospective teachers toward arithmetic. *Elem. Sch. J.*, 1951, *52*, 84-90.

77. Dutton, W. H. Measuring attitudes toward arithmetic. *Elem. School J.*, 1954, *55*, 24-31.

78. Dyer, H. S., Kalin, R. and Lord, F. M. *Problems in mathematical education*. Princeton, N. J.: Educational Testing Service, 1956.

79. Eagle, E. The relationship of certain reading abilities to success in mathematics. *Math. Teach.*, 1948, *41*, 175-179.

80. Egner, R. E., and Obelsky, A. J. The significance of stereotyped conditioning for education. *J. higher Educ.*, 1957, *28*, 329-336.

81. Eldridge, H. M. A study of the variation in accomplishment and subject preference in different secondary schools. *Dissertation Abstr.*, 1956, *16*, 1226-1227.

82. *Emerging practices in mathematics education*, Twenty-Second Yearbook of the National Council of Teachers of Mathematics. Washington, D. C.: National Council of Teachers of Mathematics, 1954.

83. Engelhard, Mary Dominic. *An experimental study of arithmetic problem-solving ability of sixth grade girls.* Washington, D. C.: The Catholic University of America Press, 1955.

84. Eysenck, H. J. Student selection by means of psychological tests - a critical survey. *Brit. J. Educ. Psychol.*, 1947, *17*, 20-39.

85. Fehr, H. F. Psychology of learning in the junior high school. *Math. Teach.*, 1956, *49*, 235-240.

86. Fehr, H. F. Teaching for appreciation of mathematics. *Sch. Sci. Math.*, 1952, *52*, 19-24.

87. Feierabend, R. L. *The role of reinforcement in the acquisition of word meanings.* Unpublished Ph.D. thesis, Yale University, 1955.

88. Flack, W. S. An investigation of mathematic ability in the classroom. *Forum of Education*, 1926, *IV*, 44-56.

89. Flournoy, Frances. Consideration of the ways children think when performing higher decade addition. *Elementary School J.*, 1957, *57*, 204-8.

90. Fouracre, L. Psychological tests of mathematical ability. *Forum of Educ.*, 1926, 201.

91. Frederiksen, N., and Satter, G. A. The construction and validation of an arithmetical computation test. *Educ. psychol. Measmt.*, 1953, *13*, 209-227.

92. Frost, N. F., and Brandes, L. G. Factors for predicting success in beginning algebra. *Calif. J. educ. Res.*, 1956, *7*, 79-83.

93. Fuller, K. G. An experimental study of two methods of long division. *Teach. Coll. contr. Educ.*, 1949, No. 951.

94. Fullerton, C. K. A comparison of the effectiveness of two prescribed methods of teaching multiplication of whole numbers. *Dissertation Abstr.*, 1955, *15*, 2126-2127.

95. Gabel, K. E. A study of the relative effectiveness of four distributions of weekly instructional time in college freshman mathematics. *Dissertation Abstr.*, 1955, *15*, 1560.

96. Gagne, R. M. Problem solving and thinking. *Annual rev. Psychology*, 1959, *10*, 147-172.

97. Garnett, P. D. Teaching arithmetic by visual training. *Optom. Wkly.*, 1956, *47*, 237-238.

98. Gebhard, Mildred E. The effect of success and failure upon the attractiveness of activities as a function of experience, expectation and need. *J. exp. Psychol.*, 1948, *38*, 371-388.

99. Gebhard, Mildred E. Changes in the attractiveness of activities: the effect of expectation preceding performance. *J. exp. Psychol.*, 1949, *39*, 404-413.

100. Georgiana, D. W. S. The relationship to two experimental measures of student motivation to academic success in college. *Dissertation Abstr.*, 1955, *15*, 1556-1557.

101. Gibb, E. G. A review of decade of experimental studies which compared methods of teaching arithmetic. *J. educ. Res.*, 1953, *46*, 603-608.

102. Gibb, E. G. Children's thinking in the process of subtraction. *J. exper. Educ.*, 1956, *25*, 71-80.

103. Goldman, Morton. Alternative classroom standards concerning management of hostility and effects of student learning. *Dissertation Abstr.*, 1955, *15*, 2098-2099.

104. Gouph, Sister Mary Tides. Mathemaphobia: causes and treatments. *Clearing House*, 1954, *28*, 290-294.

105. Grime, H. E. Aptitude and ability in elementary algebra. *Sch. Sci. Math.*, 1947, *47*, 781-784.

106. Guetzkow, H., Kelly, E. L., and McKeachie, W. J. An experimental comparison of recitation, discussion, and tutorial methods in college teaching. *J. educ. Psychol.*, 1954, *45*, 193-207.

107. Guiler, W. L. Difficulties encountered in percentage by college freshmen. *J. educ. Res.*, 1946, *40*, 81-95.

108. Guiler, W. S. Difficulties encountered by college freshmen in decimals. *J. educ. Res.*, 1946, *40*, 1-13.

109. Gunderson, Agnes C. Thought patterns of young children in learning multiplication and division. *Elem. Sch. J.*, 1955, *55*, 453-61.

110. Hadamard, J. *The psychology of invention in the mathematical field*. Princeton, N. J.: Princeton University Press, 1949.

111. Hagelberg, R. R. A study of the effectiveness of specific procedures in solving verbal arithmetic problems. *Dissertation Abstr.*, 1957, *17*, 2878-2879.

112. Hall, J. V. Solving verbal arithmetic problems without pencil or paper. *Elem. Sch. J.*, 1947, *48*, 212-217.

113. Hamza, M. Retardation in mathematics amongst grammar school pupils. *Brit. J. educ. Psychol.*, 1952, *22*, 189-195.

114. Hannon, H. The mastery of certain aspects of mathematics for general education by college students. *J. educ. Res.*, 1957, *50*, 363-71.

115. Hartung, M. L., and Fawcett, H. P. The measurement of understanding in secondary school mathematics. *Yearbook nat. soc. study Educ.*, 1946, *45*, 157-174.

116. Haslerud, G. M., and Meyers, Shirley. The transfer value of given and individually derived principles. *J. edu. Psych.*, 1958, *49*, 293-98.

117. Heidbreder, Edna. The attainment of concepts: VI. Exploratory experiments on conceptualization at perceptual levels. *J. Psychol.*, 1948, *26*, 193-216.

118. Heidbreder, Edna. The attainment of concepts: VIII. The conceptualization of verbally indicated instances. *J. Psychol.*, 1949, *27*, 263-309.

119. Helson, H., and Helson, H. B. Some common features of concrete and abstract thinking. *Amer. J. Psychol.*, 1946, *59*, 468-472.

120. Henderson, K. B. Use of resource units in teaching mathematics. *Sch. Sci. & Math.*, 1949, *49*, 345-49.

121. Hendrix, Gertrude. A new clue to transfer of training. *Elem. Sch. J.*, 1947, *48*, 197-208.

122. Henry, Jules. Docility, or giving teacher what she wants. *J. soc. Issues*, 1955, *11*(2), 33-41.

123. Henry, L. K. The role of insight in plane geometry. *J. educ. Psychol.*, 1934, *25*, 598-610.

124. Hensell, K. C. The content of problems of arithmetic. *Dissertation Abstr.*, 1956, *16*, 1857.

125. Hilgard, E. R., Irvine, R. P., and Whipple, J. E. Rote memorization, understanding, and transfer: an extension of Katona's card-trick experiments. *J. exp. Psychol.*, 1953, *46*, 4.

126. Holmes, D. C., and Harvey, J. F. An evaluation of two methods of grouping. *Educ. res. Bulletin*, 1956, *35*, 213-222.

127. Holmes, W. F. The relationship between numerical-verbal ability and educational and vocational interests. *Microfilm Abstr.*, 1949, *9*(2), 80-81.

128. Ilg, Frances L., and Ames, Louise B. Developmental trends in arithmetic. *J. genet. Psychol.*, 1951, *79*, 3-28.

129. Illoff, L. B. The effect of certain teaching practices involving systematic home-school cooperation upon the achievement of eighth-grade pupils in mathematics. *Dissertation Abstr.*, 1957, *17*, 2935.

130. Isard, Eleanore S. The development of a forced-choice inventory of attitudes toward school experience for predicting scholas[illegible] achievement. *Dissertation Abstr.*, 1955, *15*, 1436.

131. Jackson, W. N. The role of algebra in the development of relational thinking. *Dissertation Abstr.*, 1957, *52*, 2936-2938.

132. Johnson, J. T. On the nature of problem-solving in arithmetic. *J. educ. Res.*, 1949, *43*, 100-115.

133. Johnson, D. Implications of research in the psychology of learning for science and mathematics teaching. *Rev. of ed. Res.*, 1957, *27*, 400-413.

134. Kasper, Abigail, A. A study of the relationships among classroom climate, emotional adjustment, and reading achievement. *Dissertation Abstr.*, 1956, *16*, 1399-1400.

135. Kidd, K. P. Improving the learning of math. *Math. Teach.*, 1954, *47*, 393-400.

136. Kinzer, J. R., and Kinzer, Lydia G. Predicting grades in advanced college mathematics. *J. appl. Psychol.*, 1953, *37*, 182-184.

137. Koenker, R. H. Arithmetic readiness at the kindergarten level. *J. educ. Res.*, 1948, *42*, 218-223.

138. Kosofsky, S. The effectiveness of positive and negative guidance as related to the degree of the organization of the material to be learned. *J. gen. Psychol.*, 1952, *47*, 91-104.

139. Kostick, M. M. A study of transfer: sex differences in the reasoning process. *J. educ. Psychol.*, 1954, *45*, 449-58.

140. Kramer, K. A comparison of objectives, methods and achievement in arithmetic in the U.S. and in the Netherlands. *Dissertation Abstr.*, 1957, *17*, 2881.

141. Krathwohl, W. C. Relative contributions of aptitude and work habits to achievement in college mathematics. *J. educ. Psychol.*, 1953, *44*, 140-148.

142. Kress, R. A. Jr. An investigation of the relationship between concept formation and achievement in reading. *Dissertation Abstr.*, 1956, *16*, 573-574.

143. Krumboltz, J. An investigation of the effect of three teaching methods on motivational outcomes in a How to Study course. *Dissertation Abstr.*, 1955, *15*, 2470.

144. Kurtz, J.J., and Swenson, Esther J. Student, parent, and teacher attitude toward student achievement in school. *Sch. Rev.*, 1957, *59*, 273-279.

145. Kvaraceus, W. C., and Lanigan, Mary A. Pupil performance on the Iowa Every-Pupil Test of Basic Skills administered at half year intervals in the junior high school. *Educ. psychol. Measmt.*, 1948, *8*, 93–100.

146. Lankford, F. G. Jr., and Pattishall, E. G. Jr. *An experimental project in the teaching of arithmetic*. Charlottesville, Va.: University of Virginia. Council for educational research of the Richmond-Petersburg Area, 1956.

147. Laughlin, B. Frontiers in teaching mathematics and science. *Sch. Sci. & Math.*, 1951, *51*, 211–218.

148. LeBaron, W. A. A study of teachers' opinions in methods of teaching arithmetic in the elementary school. *J. educ. Res.*, 1949, *43*, 1–9.

149. *The learning of mathematics, its theory and practice*, Twenty-first Yearbook of the National Council of Teachers of Mathematics, Washington, D.C.: National Council of Teachers of Mathematics, 1953.

150. Lee, Doris M. A study of specific ability and attainment in mathematics. *Brit. J. educ. Psychol.*, 1955, *25*, 178–189.

151. Leodas, C. J. First number concepts: cornerstones of mathematics. *J. educ. Sociol.*, 1957, *30*, 343–45.

152. Love, T. A. The relation of achievement in mathematics to certain abilities in problem-solving. *Microfilm Abstr.*, 1951, *11*, 960–962.

153. Lynn, R. Tempermental characteristics related to disparity of attainment in reading and arithmetic. *Brit. J. Educ.*, 1957, *27*, 62–67.

154. MacArthur, R. S. An experimental investigation of persistence in secondary school boys. *Canad. J. Psychol.*, 1955, *9*, 42–54.

155. McAllister, B. Arithmetical concepts and the ability to do arithmetic. *Brit. J. educ. Psychol.*, 1951, *21*, 155–156.

156. McGauvran, Mary Elizabeth. A study of the relationship between attitude toward school and scholastic success at the high school and college level. *Dissertation Abstr.*, 1955, *15*, 2482–83.

157. McKenzie, Helen M. The standardization of an arithmetic-mathematics test in a Midland area. *Brit. J. educ. Psychol.*, 1952, *22*, 73–74.

158. McNemar, O. W. An attempt to differentiate between individuals with high and low reasoning ability. *Amer. J. of Psychol.*, 1955, *68*, 20–36.

159. McSwain, E. T. Discovering meanings in arithmetic. *Childh. Educ.*, 1950, *26*, 267–271.

160. Mahachek, Joy E. Growth in the uses of arithmetic concepts and processes. *Uni. of Pittsburgh, Abstracts of doctoral dissertations, 1948*. Pittsburgh, 1949, 430–439. (*Uni. Pittsburgh Bull.*, 1949, *45*, 8).

161. Mann, H. Lincoln Study Group. (Teachers Coll., Columbia U., N. Y.) Recommended: group research for teachers. *Tech. Coll. Res.*, 1948, *50*, 108–113.

162. Manning, W. H. *Correlates of change in complex functions*. Unpublished Ph.D. thesis, Washington University, January, 1959.

163. Materials and procedures to develop reading efficiency in mathematics. Ch. XII in Robinson, H. M. (ed.), *Promoting maximal reading growth among able learners*, *Suppl. ed. Mono*., 1954, *81*, 117-130.

164. Mathews, C. W. Use of mnemonic devices in mathematics. *Sch. Sci. & Math*., 1949, *49*, 491-2.

165. Mead, A. R. Illustrations of levels of learning activity: objective-improvement of thinking abilities. *Peabody J. Educ*., 1955, *33*, 107-108.

166. Mead, A. R. Transfer of training again. *J. educ. Psychol*., 1946, *37*, 390-397.

167. Mech, E. V. Performance in a verbal addition task related to pre-experimental "set" and verbal noise. *J. exp. Educ*., 1953, *22*, 1-17.

168. Meddleton, I. G. An experimental investigation into the systematic teaching of number combinations in arithmetic. *Brit. J. educ. Psychol*., 1955-56, *25-26*, 117-127.

169. Miller, M. S. An analysis of clinical cases and non-clinical cases as determined by an arithmetic proficiency test. *Dissertation Abstr*., 1957, *17*, 642-643.

170. Milton, G. A. The effects of sex-role identification upon problem-solving skill. *J. abnorm. soc. Psychol*., 1957, *55*, 2.

171. Mones, L. Psychiatric insight and educational effort. *Education*, 1955, *76*, 139-151.

172. Montessori, M. *The Advanced Montessori Method. Scientific Pedagogy as Applied to Children from Seven to Eleven Years*. *II. The Montessori Elementary Material*. New York: Stokes, 1917.

173. Morton, J. A. A study of children's mathematical interest questions as a clue to grade placement of arithmetic topics. *J. educ. Psychol*., 1946, *37*, 293-315.

174. Morton, R. L. *Teaching Arithmetic. What Research Says to the Teacher. No. 2.* Washington, D. C.: National Education Association, 1953.

175. Moser, H. E. Advancing arithmetic readiness through meaningful numbers experiences. *Childh. Educ*., 1948, *24*, 322-326.

176. Mosher, H. M. Subject preferences of girls and boys. *Sch. Rev*., 1952, *60*, 34-38.

177. Murray, J. E. An analysis of geometric ability. *J. educ. Psychol*., 1949, *40*, 118-124.

178. Nakamura, C. Y. Conformity and problem solving. *J. abnorm. soc. Psychol*., 1958, *56*, 3.

179. Norman, Martha. Three methods of teaching basic division facts. *Dissertation Abstr*., 1955, *15*, 2134.

180. Northrup, G. J., Pingry, R. E., and Winsor, A. L. A functional item analysis of group performance in general mathematics. *J. exp. Educ*., 1950, *18*, 279-285.

181. Olander, H. T., Van Wagenen, M. J., and Bishop, Helen Miriam. Predicting arithmetic achievement. *J. educ. Res*., 1949, *43*, 66-73.

182. O'Leary, M. J. The measurement and evaluation of the work habits of over-achievers and under-achievers to determine the relationship of these habits to achievement. *Dissertation Abstr*., 1955, *15*, 2104-2105.

183. Orleans, J. S., and Wandt, E. The understanding of arithmetic possessed by teachers. Elem. Sch. J., 1953, 53, 501, 512.

184. Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. The measurement of meaning. Urbana, Ill.: University of Illinois Press, 1957.

185. Ostlund, L. A. Group functioning under negative conditions. J. educ. Psychol., 1956, 47, 32-39.

186. Owens, W. A., and Johnson, Wilma C. Some measured personality traits of collegiate underachievers. J. educ. Psychol., 1949, 40, 41-46.

187. Peel, E. A. Practice effects between three consecutive tests of intelligence. Brit. J. educ. Psychol., 1952, 22, 196-199.

188. Piaget, J., and Inhelder, B. Diagnosis of mental operations and theory of the intelligence. Amer. J. ment. Def., 1947, 51, 401-406.

189. Piaget, J. How children form mathematical concepts. Scientific American, Nov. 1953, 189, 74-79.

190. Pincus, M. An investigation into the effectiveness of two methods of instruction in addition and subtraction facts. Dissertation Abstr., 1956, 16, 1415.

191. Pitts, R. J. Relationships between functional competence in mathematics and reading grade levels, mental ability, and age. J. educ. Psychol., 1952, 43, 486-492.

192. Plank, Emma N. Observations on attitudes of young children toward mathematics. Math. Teach., 1950, 43, 252-263.

193. Plank, Emma N., and Plank, R. Emotional components in arithmetical learning as seen through autobiographies. Psychoanalytic Study of the Child., 1954, 9, 274-296.

194. Plumlee, Lynette B. Comparison of problem types in the comprehensive mathematics test. Coll. Bd. Rev., 1947, 1, 17; 29-31.

195. Poffenberger, T., and Norton, D. Factors in the formation of attitudes towards mathematics, J. educ. Res., 1959, 52, 171-176.

196. Powers, S. R. (ed.) The natural sciences and mathematics. Rev. Ed. Res., 1948, 18, 293-368.

197. Purdy, C. R., and Kinney, L. B. Directing learning in arithmetic. Elem. Sch. J., 1954, 54, 285-290.

198. Rabinowitz, R. Attributes of pupils achieving beyond their level of expectancy. J. Pers., 1956, 24, 308-317.

199. Ramharter, Hazel K., and Johnson, H. C. Methods of attack used by "good" and "poor" achievers in attempting to correct errors in six types of subtraction involving fractions. J. educ. Res., 1949, 42, 586-597.

200. Reed, C. E. A study of three groups of college preparatory students who differ in relative achievement. Dissertation Abstr., 1955, 15, 2106.

201. Reeder, Thelma A. A study of some relationships between level of self-concept, academic achievement, and classroom adjustment. Dissertation Abstr., 1955, 15, 2472.

202. Reichert, E. C. The development of decimal abilities. In University of Minnesota, Summaries of Ph.D. theses, 1949, Minneapolis, Minn., 1949, 3, 81-87.

203. Reilly, J. The development of a technique to measure a component of critical thinking in the physics area. Dissertation Abstr., 1956, 16, 1415-1416.

204. Renner, J. W. Relationships between instructional provisions and functional competence in mathematics of Iowa high school seniors. Dissertation Abstr., 1955, 15, 1188.

205. Riess, Anita P. A new rationale for the teaching of fractions. Harvard educ. Rev., 1955, 25, 105-125.

206. Riess, Anita P. Number readiness in research. Chicago: Scott, Foresmam and Co., 1947.

207. Ro.nlin, I. Nonintellectual factors in the learning of general mathematics. Unpublished Ph.D. thesis, University of Chicago, 1952.

208. Rosen, V. H. On mathematical "illumination" and the mathematical thought process: a contribution to the genetic development and meta-psychology of abstract thinking. Psychoanal. Stud. Child., 1954, 8, 127-154.

209. Rusch, C. E. An analysis of arithmetic achievement in grades four, six, and eight. Dissertation Abstr., 1957, 17, 2217.

210. Sewin, E. I. Motivation in mathematics: its theoretical basis, measurement and relationships with other factors. Math. Teach., 1951, 44, 471-478.

211. Schultz, D. G., and Ricciuti, H. N. Level of aspiration measures and college achievement. J. gen. Psychol., 1954, 51, 267-275.

212. Schunert, J. The association of mathematical achievement with certain factors resident in the teacher, in the teaching, in the pupil, and in the school. J. exp. Educ., 1951, 19, 219-238.

213. Schwellenbach, J. A. An experiment in predicting the ability of eighth grade students to work simple algebra problems. Calif. J. educ. Res., 1954, 5, 36-41.

214. Semmelmeyer, Madeline. Extensional methods in dealing with abstractions in reading. Elem. Sch. J., 1949, 50, 28-36.

215. Shepard, R. Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. Psychometrika, 1957, 22, 325-345.

216. Shepler, W. D. A study of scholastic achievement in secondary school science in relation to pupils' relative preference for this subject. Dissertation Abstr., 1956, 16, 1376-1377.

217. Shimmon, W. J. The identification of some of the concomitants of learning effectiveness. Dissertation Abstr., 1953, 13, 1252-1253.

218. Sigel, I. E. Developmental trends in abstraction ability of children. Child. Develpm., 1953, 24, 131-144.

219. Signori, E. Personality, psychotherapy and critical thinking in education. Education, 1955, 76, 123-128.

220. Skinner, B. F. The science of learning and the art of teaching. Harvard Educ. Rev., Spring 1954, 24, 86-97.

221. Slinger, G. E. Freshman problems and academic achievement. Dissertation Abstr., 1955, 15, 2474.

222. Sobel, M. A. A comparison of two methods of teaching certain topics in ninth grade algebra. Dissertation Abstr., 1954, 14, 1647.

223. Sobel, M. A. Concept learning in algebra. Math. Teach., 1956, 49, 425-430.

224. Stephens, J. M., and Lichtenstein, A. Factors associated with success in teaching grade five arithmetic. J. educ. Res., 1947, 40, 683-694.

225. Stern, Catherine. *Children discover arithmetic*. New York: Harper, 1949.

226. Stokes, M. S. Instructional techniques and devices: the need and purpose. *Educ. Adm. Superv.*, 1955, *41*, 361-363.

227. Sueltz, B. A., Boynton, H., and Sauble, I. The measurement of understanding in elementary school mathematics. *Yearb. nat. Soc. Stud. Educ.*, 1946, *45*, 138-156.

228. *A Survey of mathematical education.* Research Bulletin No. 55-27, Educational Testing Service, Princeton, N. J., 1955, Multilithed.

229. Sutherland, J. A comparison of pupils' arithmetical ability in the secondary school with their ability at the time of transfer from primary schools. *Brit. J. educ. Psychol.*, 1951, *21*, 3-8.

230. Sutherland, J. An investigation into some aspects of problem solving in arithmetic. *Brit. J. educ. Psychol.*, 1941, *11*, 215-222.

231. Sutherland, J. An investigation into the prognostic value of certain arithmetic tests at the age of eleven plus. *Brit. J. Psychol. Statist. Sect.*, 1952, *5*, 189-196.

232. Sweeney, E. J. *Sex differences in problem solving*. Tech. Report I, Office of Naval Research, 1953.

233. Swenson, Esther. Organization and generalization as factors in learning, transfer, and retroactive inhibition. In Swenson, E. J., Anderson, G. L., and Stacey, C. L. *Learning theory in school situations*. Minneapolis: U. of Minnesota Press, 1949.

234. Swineford, Frances. A number factor. *J. educ. Psychol.*, 1949, *40*, 157-167.

235. Swineford, Frances. General, verbal, and spatial bi-factors after three years. *J. educ. Psychol.*, 1949, *40*, 353-360.

236. Symonds, P. M. What education has to learn from psychology. *Teachers College Record*, 1955, *56*, 277-285.

237. Taylor, D. W., and McNemar O. W. Problem solving and thinking. *Annual. Rev. of Psych.*, 1955, *6*.

238. Thiele, C. L. *The contribution of generalization to the learning of the addition facts*. New York: Teachers College, Columbia Univ., Bureau of Pub., 1938.

239. Thorndike, E. L. *New methods in teaching arithmetic.* New York: Rand-McNally, 1921.

240. Thorndike, E. L. *The psychology of arithmetic*. New York: MacMillan, 1922.

241. Thorndike, E. L. *The psychology of algebra*. New York: MacMillan, 1923.

242. Tibbetts, J. R. The role of parent-child relationships in the achievement of high school pupils: a study of the family relationships associated with under-achievement and high achievement of high school pupils. *Dissertation Abstr.*, 1955, *15*, 232.

243. Tilton, J. W. An experimental effort to change the achievement test profile. *J. exp. Educ.*, 1947, *15*, 318-323.

244. Toops, H. A. Psychological notions expressed in mathematical form. *Sch. Sci. Math.*, 1933, *33*, 479-489.

245. Traxler, A. E. Reliability and validity of a brief survey of arithmetic skills, *Educ. Rec. Bull.*, 1954, No. 62, 76-84.

246. Troyer, M. E. How does marking on the basis of ability affect learning and interests? *Amer. Psychologist*, 1948, *3*, 297, (abstract).

247. Trumbull, R. *Analysis of basic factors constituting necessary mathematical proficiency required for success in Naval Aviation: Report I. Addition*. U. S. Naval Sch. Aviat. Med. Res. Rep., 1952, Proj. NM 001 058.20.01.

248. Trumbull, R. *Analysis of basic factors constituting necessary mathematical proficiency required for success in Naval Aviation. Report II. Subtraction*. U. S. Naval Sch. Aviat. Med. Res. Rep., 1952, Proj. No. NM 001058.20.02.

249. Van Engen, H. The child's introduction to arithmetic reasoning. *Sch. Sci. Math.*, 1955, *55*, 358-363.

250. Van Engen, H. Place-value and the number system. *Suppl. educ. Monogr.*, 1947, *63*, 59-75.

251. Van Engen, H. An analysis of meaning in arithmetic, I and II. *Elem. Sch. J.*, 1949, *49*, 321-329; 395-400.

252. Wallace, D. G. A comparative analysis of achievement in three and five hour sections of freshman English. *J. educ. Res.*, 1956, *49*, 505-513.

253. Ward, J. N. Group-study versus lecture-demonstration in physical science instruction for general education in college students. *J. exp. Educ.*, 1956, *24*, 197-210.

254. Washburne, C. W. When shall we teach arithmetic? – a committee of seven investigation. *Elem Sch. J.*, 1928, *28*, 659-665.

255. Watts, Kathleen P. The effect of a university training in mathematics on the scores in an intelligence test. *Brit. J. educ. Psychol.*, 1954, *24*, 32-37.

256. Weaver, J. F. Misconceptions about rationalizations in arithmetic. *Math. Teach.*, 1951, *44*, 377-381.

257. Weaver, J. F. A crucial aspect of meaningful arithmetic instruction. *Math. Teach.*, 1950, *43*, 112-117.

258. Weitz, J., and Flurry, Christine. A study of concept formation. *Amer. Psychologist.*, 1948, *3*, 333.

259. Wertheimer, M. *Productive thinking*. New York: Harper, 1945.

260. Wilson, G. M. The social utility theory as applied to arithmetic, its research basis, and some of its implications. *J. educ. Res.*, 1948, *41*, 321-337.

261. Wrightstone, J. W. Influence of research on instruction in arithmetic. *Math. Teach.*, 1952, *45*, 187-192.

262. Young, Florence M. Causes for loss of interest in high school subjects as reported by 651 college students. *J. educ. Res.*, 1932, *25*, 110-115.

263. Youtz, R. P. The relation between number of confirmations of one hypothesis and the speed of accepting a new and incompatible hypothesis. *Amer. Psychologist*, 1948, *3*, 248-249.

Selected Foreign References

264. American Educational Research Association. Educational research in countries other than the U. S. A. *Rev. Ed. Research*, 1957, *27*, No. 1.

265. (Anon.) Arithmetical problems. *Rev. Psicol. Pedag. apl., Valencia*, 1951, *2*(3), 51-58.

266. Arshavsky, Sh. The basis of mathematical thinking. *Hahinukh*, 1950/52, *24*, 149-159.

267. Beezhold, F. W. The use of the second language as a medium for a mathematical achievement and a scientific information test. *Bull. Nat. Inst. pers. Res.*, Johannesburg, 1951, *3*(3), 9-18.

268. Bertran-Salieti, J. Main errors incurred in fundamental arithmetic operations. *Rev. Psicol. Pedag. apl., Valencia*, 1954, *5*, 93-116.

269. Biesheuvel, S. The role of arithmetical items in tests of general intelligence. *Proc. S. Afr. Psychol. Ass.*, 1951, *2*, 17-18.

270. Dabout, M., Gille, I., Joly, M., Lepez, R., and Vincent, A. Comprehension and reasoning in arithmetical problems in pupils from 11 to 13 years of age. *Enfance*, 1952, *5*, 511-522.

271. Fettweis, E. Points of controversy in the history of arithmetic in the light of ethnology and psychology. *Scientia*, 1953, *88*(497), 235-249.

272. Holtzhausen, J. A. Remedial teaching of pupils having difficulties in arithmetic. *J. soc. Res., Pretoria*, 1950, *1*, 183-194.

273. Hotyat, F. The progress of thought in the mathematical reasoning of the adolescent. *Rev. Sci. Pedag.*, 1948, *10*, 97-116.

274. Lal, M. S. Arithmetic ability in relation to socioeconomic status of students and efficiency of school instruction in general. *Indian J. Psychol.*, 1949, *24*, 66-77.

275. Lankov, A. V. *On the history of the development of progressive ideas in Russian methods (of teaching) mathematics*. Moscow: Uchpedgiz, 1951.

276. Mannoury, G. *The psycho-linguist foundations of mathematics*. Neuchatel, Switzerland: Editions du Griffon, 1947.

277. Michaud, E. The child and geometric figures. *J. Psychol. norm. path.*, 1947, *40*, 154-168.

278. Mukherji, N. Choice of curricular subject (mathematics). *Indian J. Psychol.*, 1951, *26*, 127-131.

279. Olckers, P. J. A factorial study of arithmetical ability. *J. soc. Res., Pretoria*, 1951, *2*, 1-21.

280. Piaget, J. The subtraction of partial congruent surfaces from two equal surfaces. In *Miscellanea psychologica, Albert Michotte*. Paris: Library Philosophique, 1947, 167-180.

281. Sen, J. M. Training in mathematics vocabulary. *Indian J. Psychol.*, 1949, *24*, 62-65.

282. Subes, J., and Subes, M. Tests of educational level-arithmetic and spelling. *Rev. Psychol. appl.*, 1952, *2*, 111-150.

283. Vandevelde, R. Experimental study of the problems of arithmetic. *Rev. Sci. Pedag.*, 1948, *10*, 450-454.

284. Violette, R. Researchers of school psychologists in the psycho-pedagogy of mathematics. *Enfance*, 1952, *5*, 491-510.

285. Weber, H. An investigation into the factorial structure of numerical tasks. *Z. exp. angewand. Psychol.*, 1953, *1*, 336-393.

CONFERENCE PROGRAM

Psychological Problems and Research Methods in Mathematics Training

May 1 and 2, 1959

Sponsored by Washington University under Cooperative Research Project No. 642 of the Office of Education, U. S. Department of Health, Education and Welfare

Eads Hall, Room 203, Washington University Campus

SESSION I - Friday Morning, May 1, 1959

EXPLORING THE PROBLEM AREA, Philip H. DuBois, Chairman

9:00 "Welcome," Dean Thomas S. Hall

9:10 "Opening Remarks," Randall M. Whaley

9:20 "Perspectives on Mathematics Training from a Review of the Educational and Psychological Literature," Rosalind L. Feierabend

9:30 "An Analysis of Teaching-Learning Functions," Ray C. Carpenter

10:30 "Implications of Some Doctrines of Mathematics Teaching for Research in Human Learning," Robert Gagné

11:15 "The Psychological Study of Problem-Solving," Donald W. Taylor

SESSION II - Friday Afternoon, May 1, 1959

MATHEMATICS IN THE CLASSROOM, E. G. Begle, Chairman

1:15 "The Discovery Method in Training in Arithmetic: A Demonstration with Discussion of Implications," David A. Page

2:45 "The Role of Nonverbal Awareness in the Learning of Mathematics," Gertrude Hendrix

3:30 "Teaching the Strategy of Problem-Solving," Paul C. Rosenbloom

4:15 "A Rationale for Direct Observation in Mathematics Classrooms," E. Muriel J. Wright

SESSION III - Saturday Morning, May 2, 1959

THE MEASUREMENT APPROACH, Lee J. Cronbach, Chairman

9:00 "A Method of Determining the Correlates of Change in Skill in Mathematics," Philip H. DuBois

9:30 "Scaling Methods Appropriate for Developing Measurements of Achievement for Research on Mathematics Training," Goldine C. Gleser

10:30 "Mental Abilities and Personality Variables in Relation to Proficiency in Mathematics," Saul B. Sells

11:15 "The Interaction of Personality Variables with Teaching Methods and Their Effect on Learning at the College Level," Wilbert J. McKeachie

SESSION IV - Saturday Afternoon, May 2, 1959

THE EXPERIMENTAL APPROACH, Marion E. Bunch, Chairman

1:50 "Some Perceptual Factors in Learning Mathematics," James M. Vanderplas

2:30 "Some Implications of the Experimental Analysis of Behavior for Research in Teaching," Thomas F. Gilbert

3:10 "Short Cuts and Analogue Devices in Attempts to Teach the Mathematical Transform Concepts of Feed-back Theory to Graduate Students in Psychology," George L. Kreezer

3:50 Further discussion of research designs

PARTICIPANTS

Conference on Psychological Problems and Research Methods in Mathematics Training

May 1 and 2, 1959

Sponsored by Washington University under Cooperative Research Project No. 642 of the Office of Education, U. S. Department of Health, Education and Welfare

Dr. E. G. Begle	Yale University
Dr. Marion E. Bunch	Washington University
Dr. C. Ray Carpenter	The Pennsylvania State University
Dr. David L. Clark	U. S. Office of Education
Dr. Lee J. Cronbach	University of Illinois
Dr. Philip H. DuBois	Washington University
Dr. Rosalind L. Feierabend	Washington University
Dr. Robert M. Gagné	Princeton University
Dr. Thomas F. Gilbert	Harvard University
Dr. Goldine C. Gleser	University of Illinois University of Cincinnati Medical School
Dean Thomas S. Hall	College of Liberal Arts Washington University
Miss Gertrude Hendrix	University of Illinois School Mathematics Project
Dr. George L. Kreezer	Washington University
Dr. Holbrook M. MacNeille	Washington University
Dr. Winton H. Manning	Washington University
Dr. W. J. McKeachie	University of Michigan
Professor David A. Page	University of Illinois Arithmetic Project
Dr. Paul C. Rosenbloom	University of Minnesota
Dr. Saul B. Sells	Texas Christian University
Dr. Donald W. Taylor	Yale University
Dr. James M. Vanderplas	Washington University
Dr. R. M. Whaley	Executive Director, Advisory Board on Education National Academy of Science
Dr. E. Muriel J. Wright	Washington University